

WEST Search History

DATE: Tuesday, September 16, 2003

<u>Set Name</u> side by side	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u> result set
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=ADJ</i>			
L30	L29 AND human-growth-hormone	0	L30
L29	L28 AND human insulin	236	L29
L28	L27 NOT chimera	1018	L28
L27	L26 AND fusion protein	2521	L27
L26	L24 NOT Rosen-Craig.IN.	2900	L26
L25	L24 NOT Rosen-Craig.IN/	2901	L25
L24	L23 NOT Ashkenazi-Avi.IN.	2901	L24
L23	L22 NOT Rosen-Craig-A.IN.	2971	L23
L22	L17 NOT Ashkenazi-Avi-J.IN.	3165	L22
L21	L20 AND human insulin	124	L21
L20	L19 AND chimeric protein	321	L20
L19	L18 AND fusion protein	1654	L19
L18	L17 AND ((530/300 530/350)!.CCLS.)	1689	L18
L17	L16 AND L10	3281	L17
L16	L15 AND human growth hormone	5184	L16
L15	insulin	45978	L15
L14	L13 AND hGH	119	L14
L13	L12 AND human insulin	664	L13
L12	L11 AND insulin	3281	L12
L11	L10 AND human growth hormone	4467	L11
L10	(chimeric OR fusion protein OR chimera)	49651	L10
L9	L8 AND hGH	143	L9
L8	L2 AND insulin	4231	L8
L7	L2 AND insulin AND human-growth-hormone	0	L7
L6	((C07K/)!.IPC.)	8	L6
L5	(C07K)	63212	L5
L4	((C07K/5)!.CCLS.)	0	L4
L3	C07K ((5/00)!.CCLS.)	0	L3
L2	((530/300 530/350)!.CCLS.)	13606	L2
L1	(530.CCLS.)	0	L1

END OF SEARCH HISTORY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:44:15 : Search time 34.4229 seconds
(without alignments)
225,942 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260
Sequence: 1 MPTIPLSRLEFNAPRAHR.....QPFSEAVIPKECKYSLQNP 49

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 1107863 seqs, 156726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1961.DAT.*
- 3: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL/qcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	I	Description
1	260	100.0	49	20	AA1985
2	260	100.0	92	20	AA1986
3	260	100.0	107	20	AA1987
4	260	100.0	134	20	AA1988
5	260	100.0	140	20	AA1989
6	260	100.0	150	20	AA1990
7	260	100.0	192	10	AA1991
8	260	100.0	192	10	AA1992
9	260	100.0	261	10	AA1993

10	260	100.0	262	12	AA1994
11	260	100.0	262	12	AA1995
12	260	100.0	310	11	AA1996
13	257	98.8	144	11	AA1997
14	256	98.5	204	23	AA1998
15	255	98.1	234	9	AA1999
16	255	98.1	175	23	AA2000
17	255	98.1	191	7	AA2001
18	255	98.1	191	18	AA2002
19	255	98.1	191	19	AA2003
20	255	98.1	191	20	AA2004
21	255	98.1	191	20	AA2005
22	255	98.1	191	20	AA2006
23	255	98.1	191	21	AA2007
24	255	98.1	191	22	AA2008
25	255	98.1	191	22	AA2009
26	255	98.1	191	23	AA2010
27	255	98.1	191	23	AA2011
28	255	98.1	191	23	AA2012
29	255	98.1	191	23	AA2013
30	255	98.1	191	23	AA2014
31	255	98.1	191	23	AA2015
32	255	98.1	191	23	AA2016
33	255	98.1	191	23	AA2017
34	255	98.1	191	23	AA2018
35	255	98.1	191	23	AA2019
36	255	98.1	191	23	AA2020
37	255	98.1	191	23	AA2021
38	255	98.1	191	23	AA2022
39	255	98.1	191	23	AA2023
40	255	98.1	191	23	AA2024
41	255	98.1	191	23	AA2025
42	255	98.1	191	23	AA2026
43	255	98.1	191	23	AA2027
44	255	98.1	191	23	AA2028
45	255	98.1	191	23	AA2029

ALIGNMENTS

RESULT 1
AA192855
ID: AA192855 standard; protein: 49 AA.
XX AC
AA192855:
UT 19-JAN-2000 (first entry)
XX XX
DE Human growth hormone (hGH) N-terminal fragment #1.
XX
KW Growth hormone; chaperone; intramolecular; insulin; precursor;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX OS Homo sapiens.
XX XX
FN WP09550302-A1.
XX XX
PT 07-OCT-1999.
XX XX
FF 31-MAR-1998; 98W0-CN000052.
XX XX
PR 31-MAR-1998; 98W0-CN000052.
XX XX
(TONG) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX Can 2:
XX
DB WP1: 1999-610839/52.
XX
PT New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin

CC protein, which contains a thrombin recognition site, and human beta nerve
 CC growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to
 CC control geriatric dementia and other nervous disorders, and can be
 CC released from the fusion protein by incubation with thrombin (see
 CC AAN90577-8, AAP91034, AAP91299).

XX SQ Sequence 140 AA;
 Query Match 100.0%; Score 260; DB 10; Length 140;
 Best Local Similarity 100.0%; Pred. No. 7.3e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 |||||
 DB 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 6
 AAY42861
 ID AAY42861 standard; protein: 150 AA.

XX AC AAY42861;

XX DT 19-JAN-2000 (first entry)

XX DE Chimeric protein, SEQ ID 7.

XX KW Insulin; precursor; growth hormone; chaperone; intramolecular
 KW folding; conformation; chimeric protein; cleavable; recombinant;
 KW production; yield.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN W09950302-A1.

XX PD 07-OCT-1999.

XX PF 31-MAR-1998; 98WO-CN00352.

XX PR 31-MAR-1998; 98WO-CN00352.

XX PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX PI Can 2;

XX PF 1999-6i0839/52.

XX PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.

XX PS Claim 14; Page 30-31; 46pp; English.

XX CC This sequence represents a chimeric protein, which contains an
 CC N-terminal fragment of human growth hormone (hGH) of the sequence given
 CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
 CC precursor comprising insulin A and B chains (AAY42859). The hGH portion
 CC of the chimeric protein acts as an intramolecular chaperone (IMC) for
 CC the insulin precursor, enabling it to fold correctly. The cleavable
 CC peptide linker has a C-terminal Arg residue which enables the hGH portion
 CC of the chimeric protein to be removed after folding has taken place.
 CC Production of recombinant human insulin via an hGH-proinsulin chimeric
 CC protein can provide human insulin with correctly linked cysteine bridges
 CC with fewer necessary procedural steps, and hence resulting in a higher
 CC yield of human insulin. The IMC sequences not only protect insulin
 CC sequences from intracellular degradation by a microorganism host, but
 CC also promote the folding of the fused insulin precursor, facilitate the
 CC solubility of the fusion protein and decrease the intermolecular
 CC interactions among the fusion proteins, thus allowing folding of the
 CC fused insulin precursor at commercially useful high concentrations. The
 CC procedural steps of cyanogen bromide cleavage, oxidative sulphatolysis
 CC and related purification steps can thus be eliminated, along with the use
 CC of high concentrations of mercaptan or the use of hydrophobic absorbent

CC Iesins.
 XX SQ Sequence 150 AA;

Query Match 100.0%; Score 260; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 7.9e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 |||||
 DB 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 7

AAP90129

ID AAP90129 standard; protein: 192 AA.

XX AC AAP90129;

XX DT 25-MAR-2003 (updated)

XX DT 06-FEB-1996 (revised)

XX DT 01-NOV-1989 (first entry)

XX DE Human growth hormone.

XX DE Human growth hormone; fusion protein; recombinant

XX KW vector.

XX OS Homo sapiens (Human).

XX PN JPC0144981-A.

XX PD 07-JUN-1989.

XX PF 02-DEC-1987; 87JP-0304937.

XX PR 02-DEC-1987; 87JP-0404937.

XX PA (WAKT) WAKUNAGA SEIYAKU KK.

XX PF 1989-209284/29.

XX PR N-PSDB; AAN90265.

XX PT Recombinant vector contg. fusion protein - consisting of human
 PT growth hormone or deriv. ligated to foreign protein, for stability
 PT and high yield.

XX PS Disclosure; Fig 1; 19pp; Japanese.

XX CC The invention consists of a vector contg. a fusion protein which is
 CC formed by ligating, downstream of a promoter, hGH or a deriv. (pref.
 CC formed by substitn. of Met-14 with Leu) and a foreign protein.
 CC Stability of the vector in the host is greatly increased so the
 CC protein yield is higher.
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 192 AA;

Query Match 100.0%; Score 260; DB 10; Length 192;
 Best Local Similarity 100.0%; Pred. No. 1e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 |||||
 DB 1 MFPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 8

AAW92264

ID AAW92264 standard; protein: 192 AA.

XX AC AAW92264;

XX XX

DI 08-JUN-1999 (first entry)

DE Human anti-angiogenic peptide hGH Met-1-phosphat.

XX Human: anti-angiogenic; prolactin; placental lactogen; hGH; anti-angiogenic; growth hormone; hGH; hGH-V; capillary endothelial cell proliferation; placental vascularisation; pregnancy; treatment; angiogenic disease; tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation; arthritis; atherosclerotic plaques; corneal graft neovascularisation; wound healing; proliferative retinopathy; macular degeneration; fracture; psoriasis; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome; ulcer; leukaemia; reproductive disorder; contraceptive agent; gene therapy; pre-eclampsia; intrauterine growth retardation; placental dysfunction.

OS Homo sapiens.

XX W09851323-AL.

XX 19-NOV-1993.

XX 12-MAY-1993; 98WO-US9599.

XX 13-MAY-1997; 97US-0046394.

XX (REGC) UNIV CALIFORNIA.

XX Martini JA, Struman I, Taylor R, Weiner R1;

XX WPI: 1999-045192/04.

XX N-PSDB; AAX01706.

XX New anti-angiogenic peptides - comprise N-terminal fragments of human placental lactogen, human growth hormone, growth hormone variant or human prolactin

XX Example 3: Page 49; 87pp; English.

XX This invention describes novel human anti-angiogenic peptides derived from 10 to 150 consecutive amino acids selected from the N-terminal end of human placental lactogen (hPL), human growth hormone (hGH), growth hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit capillary endothelial cell proliferation and organisation (ii) inhibit angiogenesis in chick chorioallantoic membrane and (iii) binds to at least one specific receptor which does not bind an intact full length hGH, hPL, prolactin or hGH-V. The invention also describes a method for diagnosing a probable abnormality of placental vascularisation during pregnancy. The peptides can be used for treating an angiogenic disease in a subject, for inhibiting tumour formation or growth in a patient, for modulating vascularisation of a patient's placenta. In particular, the peptides can be used for preventing or treating e.g. malignant tumours, angiofibroma, arteriovenous malformation, arthritic such as rheumatoid arthritis, atherosclerotic plaques, corneal graft neovascularisation, delayed wound healing, proliferative retinopathy such as diabetic retinopathy, macular degeneration, granulations such as those occurring in hemophilic joints, inappropriate vascularisation in wound healing such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis, pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours, Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose veins, leukaemia, and reproductive disorders such as follicular and luteal cysts and choriocarcinoma. They can also be used as contraceptive agents. DNA encoding the peptides can be used in gene therapy. The measurement of abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL can be used in assays for impairment of vascular development associated with pre-eclampsia, intrauterine growth retardation, and placental dysfunction.

XX Sequence 192 AA;

XX Query Match 100.0%; Score 250; DB 20; Length 192;

XX Best Local Similarity 100.0%; Pred. No. 1e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTIPISRLFDNAMLKAHLHQLAFTYQEFPEAYIPKEQKYSFLQNP 49

DB 1 MEPTIPLSRFDNAMLKAHLHQLAFTYQEFPEAYIPKEQKYSFLQNP 49

RESULT 9

AA91299

XX AA91299 standard; protein: 261 AA.

XX AAP91299;

XX 14-DEC-1989 (first entry).

XX Human nerve growth factor and human growth hormone fusion protein.

XX Human nerve growth factor; fusion protein; thrombin;

XX geriatric dementia; nervous disorders; human growth hormone.

XX Homo sapiens (human).

XX Key Location/Qualifiers

XX Region 1..146

XX Region 141..143

XX Region 144..261

XX EP329175-A.

XX 23-AUG-1989.

XX 17-FEB-1989; 89EP-0102795.

XX 19-FEB-1988; 88JP-0035042.

XX (TOYJ) TOSOH CORP.

XX Ohtsuka E;

XX WPI: 1989-243092/34.

XX New human nerve growth factor gene encoding fusion protein

XX - having cleavage site for thrombin, useful for treating geriatric dementia, etc.

XX Claim 36; page 31-32; 38pp; English.

XX Fusion protein consisting of human growth hormone at the N-terminal end (1st region), a 3 amino acid sequence representing thrombin recognition site, and human beta nerve growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to control geriatric dementia and other nervous disorders, and can be released from the fusion protein by incubation with thrombin (see AAN90577-8, AAP91034, AAP91041).

XX Sequence 261 AA;

XX Query Match 100.0%; Score 250; DB 10; Length 261;

XX Best Local Similarity 100.0%; Pred. No. 1.5e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTIPISRLFDNAMLKAHLHQLAFTYQEFPEAYIPKEQKYSFLQNP 49

DB 1 MEPTIPLSRFDNAMLKAHLHQLAFTYQEFPEAYIPKEQKYSFLQNP 49

RESULT 10

AA91033

XX AAP91033 standard; Protein: 262 AA.

XX AAP61033;

XX 25-OCT-1991 (first entry)

XX Human beta-nerve growth factor gene product.
 XX
 KW Beta-NGF; E.coli; ds.
 XX

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Protein 145..262

XX JP61205465-A.

XX 11-SEP-1986.

XX 09-MAR-1985; 85JP-0045773.

XX 09-MAR-1985; 85JP-0045773.

XX (OTSU/; OTSUKA E.

XX WPI: 1986-281696/43.

XX Gene segment of human nerve growth factor - used in prodn. of
 PT NGF-producing recombinant Escherichia strain.

XX Claim 32; Page 482; 71pp; Japanese.

XX The protein is a direct translation of the upstream tryptophan
 CC promoter-operator lacking its attenuation sequence and human
 CC beta-NGF sequence. The product may be efficiently expressed from a
 CC transformed E.coli expression system.
 CC See also AAN63816-7.

XX Sequence 262 AA:

Query Match 100.0%; Score 260; DB 7; Length 262;
 Best Local Similarity 100.0%; Pred. No. 1.5e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

RESULT 11

AAR11740

ID AAR11740 standard; Protein; 262 AA.

XX AAR11740:

XX 25-MAR-2003 (updated)

DT 25-JUN-1991 (first entry)

XX Human growth hormone/human nerve growth factor beta fusion protein.

XX hGH; hNGF; nervous system diseases; dementia.

XX Homo sapiens.

XX JP03067598-A.

XX 22-MAR-1991.

XX 07-AUG-1989; 89JP-0202835.

XX 07-AUG-1989; 89JP-0202835.

XX (TOYJ) TOSOH CORP.

XX WPI: 1991-128768/18.

XX N-PSDB; AAO11578.

XX Purificn. of human neuron growth factor beta-subunit-contg. protein -

PT by contacting with gel having cation exchange gp. in presence of
 PT urea

PS Disclosure : fig 1; 7pp; Japanese.

XX A recombinant human nerve growth factor beta subunit-contg. protein
 CC can be produced as this fusion protein. It is purified by contacting
 CC a gel having a cation exchange gp. with the fusion protein, in the
 CC presence of urea. The purified protein is useful in a medicament
 CC for treating disorders of the nervous system, eg dementia.
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 262 AA:

Query Match 100.0%; Score 260; DB 12; Length 262;
 Best Local Similarity 100.0%; Pred. No. 1.5e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

RESULT 12

AAR03255

ID AAR03255 standard; Protein; 310 AA.

XX AAR03255;

XX 19-JUL-1990 (first entry)

XX Fusion protein of B-cell stimulatory factor-2 and B-cell

DE differentiation factor.

XX B-cell stimulatory factor-2; interleukin-6; B-cell differentiation;
 KW interleukin-5; fusion protein.

XX Homo sapiens.

XX JP02013375-A.

XX 01-JUL-1988; 88JP-0162556.

XX 01-JUL-1988; 88JP-0162556.

XX (TOYJ) TOSOH CORP.

XX WFI: 1990-062207/09.

XX N-PSDB; AAO02028.

XX Prepn. of human B-cell differentiation factor - from specified DNA
 PT sequence segment, by recombinant DNA technique, gives protein of
 PT specified amino acid sequence.

XX Claim 31; Page 9; 17pp; Japanese.

XX The protein is produced by fusing DNA encoding BDF (IL-) with DNA
 CC encoding BSF-2 (IL-5) and ligating the product into an expression vector
 CC See also AAR05313 and AAR05313.

XX Sequence 310 AA:

Query Match 100.0%; Score 260; DB 11; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.8e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 1 MFPTIPLSKLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

```

RESULT 13
AAR05313
ID AAR05313 standard; protein: 144 AA.
XX
XX
AC AAR05313;
XX
DT 19-JUL-1990 (first entry)
XX
DE Segment of B-cell stimulatory factor-2 (IL-5).
XX
KW B-cell stimulatory factor-2; interleukin-5.
XX
OS Homo sapiens.
XX
PN JP02013375-A.
XX
PD 17-JAN-1990.
XX
PF 01-JUL-1988; 88JP-G162556.
XX
PR 01-JUL-1988; 88JP-0162556.
XX
PA (TOYO) TOSOH CORP.
XX
DR WPI: 1990-062207/03.
DR N-PSDB; AAQ02C28.
XX
PT prepn. of human B-cell differentiation factor - from specified DNA
PT sequence segment, by recombinant DNA technique, gives protein of
PT specified amino acid sequence.
XX
PS Disclosure: Page 9; 17pp; Japanese.
XX
CC The sequence encoding this protein can be fused with DNA encoding B-cell
CC differentiation factor (IL-6); and ligated into an expression vector for
CC prodn. of a fusion protein.
CC See also AAR05311.
XX
SQ Sequence 144 AA;

Query Match 98.8%; Score 257; DB 11; Length 144;
Best Local Similarity 98.0%; Pred. No. 1.5e-24;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49

RESULT 14
ABB77327
ID ABB77327 standard; Protein: 204 AA.
XX
XX
AC ABB77327;
XX
DT 17-JUN-2002 (first entry)
XX
DE Human growth hormone with collagenase recognition site.
XX
KW Human: growth hormone; collagenase; recognition site.
XX
OS Homo sapiens.
XX
PN KR289691-B.
XX
PD 15-MAY-2001.
XX
PF 28-DEC-1993; 93KR-00303-B.
XX
PR 28-DEC-1993; 93KR-00303-B.
XX
PA (GLDS) LG CHEM LTD.
XX

Query Match 98.1%; Score 255; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 3e-24;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

PI YOO CG, Song YH;
XX
DQ WPI: 2002-185396/24.
UR N-PSDB; ABL5599.
XX
PT Recombinant human growth hormone having collagenase recognition region -
XX
PS Disclosure: Fig 3; 5pp; Korean.
XX
CC The invention relates to recombinant human growth hormone having a
CC collagenase recognition region.
XX
SQ Sequence 204 AA;

Query Match 98.5%; Score 256; DB 23; Length 204;
Best Local Similarity 98.0%; Pred. No. 3.5e-24;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13 VFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 61

RESULT 15
AAP81226
ID AAP81226 standard; protein: 138 AA.
XX
AC AAP81226;
XX
DQ 25-MAR-2003 (updated)
DQ 20-NOV-1990 (first entry)
XX
DE Sequence of protein with somatomedin-like activity.
XX
KW Growth hormone.
XX
OS Synthetic.
XX
PN JP63167798-A.
XX
PD 11-JUL-1988.
XX
PF 29-DEC-1986; 86JP-0310177.
XX
PR 29-DEC-1986; 86JP-G310177.
XX
PA (TOYO) TOYO SODA MFG CO LTD.
XX
DQ WPI: 1986-232632/33.
DQ N-PSDB; AAN81605.
XX
PT Polypeptide with somatomedin-like activity -
PT by culturing bacterium transformed by plasmid contg. gene
PT segment with specified DNA sequence
XX
PS Claim 2(1); Page 609; 9pp; Japanese.
XX
CC The polypeptide (AAP81226) with somatomedin-like activity and the DNA
CC (AAN81605) encoding it are claimed. A vet. residual gp. may be added to
CC the N-terminal. The polypeptide acts on the bone structure of mammals,
CC including humans, to promote bone growth. The polypeptide has high
CC production rate and is easily extracted from bacterial culture medium
CC and refined for use as a bone growth accelerator.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 136 AA;

Query Match 98.1%; Score 255; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 3e-24;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1 FTIPLSRFDNAMLPAHRLHQALNFTYGEFFRAYIPRQKYSFLQNF 43

Search completed: September 15, 2003, 12:00:56
Job time : 35.4229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Inc.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:57:43 : Search time 11,2401 seconds
(without alignments)
184,449 Million cell updates/sec

Title: US-09-423-100-2

Perfect score: 280

Sequence: 1 MFPTPLSRFLNPMLEAHR.....QEFEEAYPKCKYKFLQNP 49

Scoring table:

BLC6M62

Gapop 10.0 , Gapext: 0.5

Searched: 328717 seqs, 42310658 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2003C00000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the results being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	192	1	US-08-093-363-1
2	255	98.1	191	4	US-09-284-878-5
3	255	98.1	194	2	US-08-383-621-4
4	255	98.1	194	3	US-08-459-956-4
5	255	98.1	217	3	US-06-589-024-10
6	255	98.1	217	3	US-08-784-582-10
7	255	98.1	217	3	US-08-785-271-19
8	255	98.1	217	3	US-08-759-528-11
9	255	98.1	217	4	US-09-284-878-1
10	255	98.1	241	4	US-09-424-620B-25
11	255	98.1	245	4	US-09-285-030-56
12	255	98.1	274	3	US-08-784-582-71
13	255	98.1	360	3	US-08-784-582-73
14	249	95.8	193	4	US-05-465-461-1
15	249	95.8	217	1	US-08-187-756C-4
16	249	95.8	217	1	US-08-469-486-5
17	249	95.8	217	2	US-08-469-658-51
18	249	95.8	217	2	US-08-710-324A-4
19	249	95.8	217	4	US-09-411-657-4
20	248	95.4	191	3	US-08-800-215C-16
21	248	95.4	191	3	US-08-800-215C-18
22	248	95.4	191	3	US-08-800-215C-20
23	248	95.4	400	4	US-09-420-819-37
24	248	95.4	401	4	US-09-420-819-36
25	237	91.2	71	1	US-08-314-586-24
26	233	89.6	70	1	US-07-920-519-24
27	233	89.6	70	3	US-08-115-753-26

28 164.5 63.3 191 1 US-08-458-824-8
29 164 63.1 176 3 US-08-791-728-1
30 164 63.1 176 4 US-08-990-774-1
31 161.5 62.1 191 1 US-07-963-331D-4
32 159.5 61.3 190 1 US-08-388-267C-2
33 159.5 61.3 190 4 US-09-277-720-2
34 159.5 61.3 191 6 5210180-1
35 159.5 61.3 193 1 US-07-621-197C-2
36 159.5 61.3 193 1 US-08-363-982-2
37 159.5 61.3 193 2 US-08-383-621-1
38 159.5 61.3 193 3 US-08-459-906-1
39 159.5 61.3 216 2 US-09-105-651-1
40 159.5 61.3 216 2 US-09-105-651-3
41 158 60.8 176 3 US-08-791-728-2
42 158 60.8 176 4 US-08-990-774-2
43 158 60.8 177 1 US-08-187-756C-6
44 158 60.8 177 2 US-08-710-324A-6
45 158 60.8 177 4 US-09-411-657-6

ALIGNMENTS

RESULT 1

US-08-093-363-1
: Sequence 1, Application US/08093383
: Patent No. 5489529
: GENERAL INFORMATION:
: APPLICANT: DeBoer, Herman A.
: APPLICANT: Heyneker, Herbert L.
: APPLICANT: Seeburg, Peter H.
: TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/093.383
: FILING DATE: 14-JUL-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/519827
: FILING DATE: 28-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/198824
: FILING DATE: 05-APR-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 06/632361
: FILING DATE: 19-JUL-1984
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 06/303687
: FILING DATE: 18-SEP-1981
: ATTORNEY/AGENT INFORMATION:
: NAME: Johnston, Sean A.
: REGISTRATION NUMBER: P35,910
: REFERENCE/DOCKET NUMBER: 46C4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-3552
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 192 amino acids
: TYPE: amino acid

Sequence 8, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 2, Appli
Patent No. 5210180
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli

APPLICANT: No. 6116707mington, Karl D.

1 STATE: Texas
2 COUNTRY: USA
3 ZIP: 77210
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patent In Release #1.0, Version #1.0
9 CURRENT APPLICATION DATA:
10 APPLICANT: US/08/785-271
11 FILING DATE: Concurrently Herewith
12 CLASSIFICATION: 435
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 08/589,028
15 FILING DATE: 19-JAN-1996
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Highlander, Steven L.
18 REGISTRATION NUMBER: 37,842
19 REFERENCE/DOCKET NUMBER: CISH:513
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: 512/474-7577
22 TELEFAX: 512/474-7577
23 INFORMATION FOR SEQ ID NO: 10:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 217 amino acids
26 TYPE: amino acid
27 STRANDEDNESS: linear
28 TOPOLOGY: linear
29 US-08-785-271-10

Query Match 98.1%; Score 255; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74

RESULT 8
US-08-759-628-11
Sequence 11, Application US/08/59628
Patent No. 6225446
GENERAL INFORMATION:
APPLICANT: Altmann, Scott W.
APPLICANT: Rock, Fernando L.
APPLICANT: Bazan, J. Fernando
APPLICANT: Kastlein, Robert A.
TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,628
FILING DATE: 05-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 56/606,574
FILING DATE: 06-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,690
REFERENCE/DOCKET NUMBER: DX05520

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Peptide
LOCATION: 32..53
FEATURE:
NAME/KEY: Peptide
LOCATION: 94..115
FEATURE:
NAME/KEY: Peptide
LOCATION: 133..153
FEATURE:
NAME/KEY: Peptide
LOCATION: 192..210
OTHER INFORMATION: /color "The peptides above are
OTHER INFORMATION: depicted in Figure 1"
US-09-423-100-1-11

Query Match 98.1%; Score 255; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74

RESULT 9
US-09-284-878-1
Sequence 1, Application US/09284878
Patent No. 6342375
GENERAL INFORMATION:
APPLICANT: Olazaran, Martha Guerrero
APPLICANT: Saldana, Hugo Barrera
APPLICANT: Salgado, Jose Maria Viader
TITLE OF INVENTION: Genetically Modified Methylothrophic P. pastoris Yeast for t
TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
FILE REFERENCE: 1829,00:0000
CURRENT APPLICATION NUMBER: US/09/284,878
CURRENT FILING DATE: 1999-07-21
PRIOR APPLICATION NUMBER: PCT/MX97/00033
PRIOR FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 217
TYPE: PPT
ORGANISM: Homo sapiens
US-09-284-878-1

Query Match 96.1%; Score 255; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74

RESULT 10
US-09-424-620B-25
Sequence 25, Application US/09424620B
Patent No. 639,585
GENERAL INFORMATION:
APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.

JANG, Ki-Ryoung
 MOON, Jae-Woong
 BAE, Cheon-Soon
 YANG, Doo-Suk
 LEE, Jae-Won
 SEONG, Baik-Lin
 TITLE OF INVENTION: Process for preparing recombinant proteins using highly efficient expression vector from *Saccharomyces cerevisiae*
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BACHMAN & LAPOINTE, P.C.
 STREET: Suite 1201, 900 Chapel Street
 CITY: New Haven
 STATE: Connecticut
 COUNTRY: U.S.A.
 ZIP: 06510-2802
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 COMPUTER: IBM
 OPERATING SYSTEM: WINDOWS 95/98
 SOFTWARE: MS WORD
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/03/424,620B
 FILING DATE: 24-NOV-1999
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 241 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: PROTEIN
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:
 US-09-424-620B-25

Query Match 98.1%; Score 255; DB 4; Length 241;
 Best Local Similarity 100.0%; Pred. No. 3.3e-29;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 DQ 51 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 93

RESULT 1:
 US-09-280-030-66
 Sequence 66, Application US/09/280030A
 Patent No. 6526595
 GENERAL INFORMATION:
 APPLICANT: SATO, Seiji
 APPLICANT: Higashikuni, Naohiko
 APPLICANT: Kudo, Toshiyuki
 APPLICANT: Kondo, Masaaki
 TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
 TITLE OF INVENTION: DNAS
 FILE REFERENCE: 382,192s
 CURRENT APPLICATION NUMBER: US/09/280,030A
 CURRENT FILING DATE: 1999-03-26
 EARLIER APPLICATION NUMBER: JP10-87339/1998
 EARLIER FILING DATE: 1998-03-11
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 66
 LENGTH: 245
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Designated is an amino acid sequence of Mafsp-Mafmp20-1EV-3-GH
 US-09-280-030-66

Query Match 98.1%; Score 255; DB 4; Length 245;
 Best Local Similarity 100.0%; Pred. No. 3.3e-29;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 DQ 55 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 102

RESULT 12
 US-08-784-582-71
 Sequence 71, Application US/08/784582
 Patent No. 6110707
 GENERAL INFORMATION:
 APPLICANT: Newgard, Christopher B.
 APPLICANT: Halban, Philippe A., Karl D.
 APPLICANT: No. 6110737mington, Karl D.
 APPLICANT: Clark, Samuel A.
 APPLICANT: Thigpen, Anice E.
 APPLICANT: Quade, Christian
 APPLICANT: Kruse, Fred
 APPLICANT: McGarry, Dennis
 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM SECRETORY CELL LINES
 NUMBER OF SEQUENCES: 79
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/784,582
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/028,427
 FILING DATE: 15-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/589,028
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.
 REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: JTSO:514
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-1000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 71:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 274 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-784-582-71

Query Match 98.1%; Score 255; DB 3; Length 274;
 Best Local Similarity 100.0%; Pred. No. 3.8e-29;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 49
 DQ 27 FPTIPSLRFLDNAMLRHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNP 74

RESULT 13
 US-08-784-582-73
 Sequence 73, Application US/08/784582
 Patent No. 6110707
 GENERAL INFORMATION:

APPLICANT: Newgard, Christopher B.
APPLICANT: Halban, Philippe A.
APPLICANT: No. 6110707:ington, Karl D.
APPLICANT: Clark, Samuel A.
APPLICANT: Thiepen, Anice E.
APPLICANT: Ouade, Christian
APPLICANT: Kruse, Fred
APPLICANT: McGarity, Dennis
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4423
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,582
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,427
FILING DATE: 15-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 68/589,028
FILING DATE: 19-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:514
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 366 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-784-582-73

Query Match 98.18; Score 255; DB 3; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.4e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74

RESULT 14
US-09-465-461-1
Sequence 1; Application US/09465461
Patent No. 634844
GENERAL INFORMATION:
APPLICANT: CHAPPEL, Scott
TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and enhance recovery
FILE REFERENCE: CHAPPEL-6.1
CURRENT APPLICATION NUMBER: US/09/465,461
CURRENT FILING DATE: 1993-12-17
PRIOR APPLICATION NUMBER: 60/112,668
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patent in version 3.1
SEQ ID NO 1

LENGTH: 191;
TYPE: PRI
ORGANISM: homo sapiens
US-09-465-461-1

Query Match 95.8%; Score 249; DB 4; Length 191;
Best Local Similarity 97.9%; Pred. No. 1.8e-28;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 48

RESULT 15
US-08-187-756C-4
Sequence 4; Application US/08187756C
Patent No. 5597709
GENERAL INFORMATION:
APPLICANT: ROSEN, El AL.
TITLE OF INVENTION: Human Growth Hormone
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNEL BAIN, GILFILLIAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/187,756C
FILING DATE: January 27, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-2700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-187-756C-4

Query Match 95.8%; Score 249; DB 1; Length 217;
Best Local Similarity 97.9%; Pred. No. 2.1e-28;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRLEFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74

Search completed: September 15, 2003, 12:05:30
Job time : 12.2401 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 12:33:35 : Search time 20.544 seconds
(without alignments)
347.945 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MFPIPLSLRFDNMLRAHR.....ORRREAYIKKCKYFLNP 49

Scoring table: BLOSUM62

Gapop 15.0, Gapext 0.5

Searched: 541936 seqs, 1459.2426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptocata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptocata/1/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptocata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 6: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 7: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 8: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 9: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 10: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 11: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 12: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 13: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 14: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 15: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 16: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 17: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
- 18: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	260	100.0	49	US-10-054-873-1	Sequence 1, Appl
2	260	100.0	92	US-10-054-873-2	Sequence 2, Appl
3	260	100.0	107	US-10-054-873-6	Sequence 6, Appl
4	260	100.0	150	US-10-054-873-7	Sequence 7, Appl
5	255	98.1	191	US-09-984-010-23	Sequence 23, Appl
6	255	98.1	191	US-10-153-207-1	Sequence 1, Appl
7	255	98.1	191	US-10-400-377-1	Sequence 1, Appl
8	255	98.1	191	US-10-400-708-1	Sequence 1, Appl
9	255	98.1	214	US-10-153-207-6	Sequence 6, Appl
10	255	98.1	217	US-09-929-218-9	Sequence 9, Appl
11	255	98.1	245	US-09-280-030-65	Sequence 66, Appl
12	249	95.8	217	US-09-853-688-2	Sequence 2, Appl
13	249	95.8	217	US-09-969-7480-4	Sequence 4, Appl
14	248	95.4	217	US-09-804-409A-16	Sequence 16, Appl
15	242	93.1	217	US-09-853-688-4	Sequence 4, Appl

16	239	92.9	191	12	US-09-824-200-12	Sequence 12, Appl
17	199	76.5	217	9	US-09-850-887-3	Sequence 3, Appl
18	197	75.8	191	15	US-10-043-487-150	Sequence 350, Appl
19	197	75.8	191	12	US-10-153-207-2	Sequence 2, Appl
20	197	75.8	229	15	US-10-153-313-411	Sequence 411, Appl
21	156	72.3	246	15	US-10-188-246-18	Sequence 18, Appl
22	159.5	61.3	190	12	US-10-153-207-3	Sequence 3, Appl
23	142	54.6	54	9	US-09-876-478-14	Sequence 14, Appl
24	114	43.8	54	9	US-09-876-478-2	Sequence 2, Appl
25	114	43.8	54	9	US-09-876-478-3	Sequence 3, Appl
26	106	40.8	28	9	US-09-876-478-4	Sequence 4, Appl
27	106	40.8	28	9	US-09-876-478-7	Sequence 7, Appl
28	106	40.8	54	9	US-09-876-478-6	Sequence 6, Appl
29	88	33.8	41	15	US-10-191-879-22	Sequence 22, Appl
30	88	33.8	188	15	US-10-191-879-19	Sequence 19, Appl
31	88	33.8	210	15	US-10-191-879-10	Sequence 10, Appl
32	87	33.5	366	10	US-09-887-569A-2	Sequence 2, Appl
33	87	33.5	388	12	US-10-322-746-4	Sequence 4, Appl
34	83	31.9	56	9	US-09-876-478-15	Sequence 15, Appl
35	80	30.8	198	15	US-10-140-293-32	Sequence 32, Appl
36	80	30.8	199	15	US-10-140-293-33	Sequence 33, Appl
37	73	28.1	46	15	US-10-100-679-61	Sequence 61, Appl
38	73	28.1	46	15	US-10-100-679-62	Sequence 62, Appl
39	73	28.1	46	15	US-10-100-679-63	Sequence 63, Appl
40	72.5	27.9	125	14	US-10-036-869-25	Sequence 25, Appl
41	72.5	27.9	253	14	US-10-036-869-27	Sequence 27, Appl
42	71	27.3	197	15	US-10-140-293-13	Sequence 13, Appl
43	70	26.9	199	15	US-10-140-293-22	Sequence 22, Appl
44	70	26.9	199	15	US-10-140-293-23	Sequence 23, Appl
45	70	26.9	199	15	US-10-140-293-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-1
: Sequence 1, Application US/10054873
: Publication No. US26020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Ru
: TITLE OF INVENTION: Intramolecular Chaperone-Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESS: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentic Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/00052
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 11-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 46,946
: REFERENCE/DOCKET NUMBER: 020167-000130US
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid

```

; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-054-873-1
Query Match      100.0%   Score 260;   DB 14;   Length 49;
Best Local Similarity 100.0%   Pred. No. 5.6e-28;
Matches 49;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||*
DB 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||

RESULT 2
US-10-054-873-2
; Sequence 2, Application US/10054873
; Publication No. US20020154712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-054-873-5
Query Match      100.0%   Score 260;   DB 14;   Length 107;
Best Local Similarity 100.0%   Pred. No. 1.4e-27;
Matches 49;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||
DB 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||

RESULT 4
US-10-054-873-7
; Sequence 7, Application US/10054873
; Publication No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-054-873-2
Query Match      100.0%   Score 260;   DB 14;   Length 92;
Best Local Similarity 100.0%   Pred. No. 1.2e-27;
Matches 49;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||
DB 1 MEPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
      *|||||*|||||*|||||*|||||*|||||*|||||*|||||*|||||

RESULT 3
US-10-054-873-6
; Sequence 6, Application US/10054873
; Publication No. US20020164712A1
; GENERAL INFORMATION:

```

```
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/CN98/00052
? FILING DATE: 31-MAR-1998
? APPLICATION NUMBER: US 09/423,100
? FILING DATE: 11-DEC-2000
? ATTORNEY/AGENCY INFORMATION:
? NAME: Mycrott, Frank J
? REGISTRATION NUMBER: 46,946
? REFERENCE/DOCKET NUMBER: 020167-0001A005
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 150 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-054-873-7

Query Match 100.0%; Score 255; DB 11; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MEPTPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 49
      ||||| ||||| || ||||| ||||| ||||| ||||| |||||
DB 1 MEPTPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 49

RESULT 5
US-09-984-010-23
? Sequence 23, Application US/09984010
? Publication No. US20030164578A1
? GENERAL INFORMATION:
? APPLICANT: Ballance, David James
? TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
      AND SERUM ALBUMIN
? NUMBER OF SEQUENCES: 26
? CORRESPONDENCE ADDRESS:
? ADDRESSER: FINNEGAN, HENDERSON, FARROW, GARRETT & DUNNEN, LLP
? STREET: 1300 I Street, NW
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20005-3314
? COMPUTER READABLE FORM:
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent; Release #1.0, Version #1.30 (EPO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/984,010
? FILING DATE: 21-May-2002
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 09/091,873
? FILING DATE: 25-JUN-1998
? APPLICATION NUMBER: PCT/GB96/04164
? FILING DATE: 19-DEC-1996
? INFORMATION FOR SEQ ID NO: 23:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 191 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-984-010-23

Query Match 98.1%; Score 255; DB 11; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

```
QY 2 FETPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 49
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FETPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 48

RESULT 6
US-10-153-207-1
? Sequence 1, Application US/10153207
? Publication No. US20030153063A1
? GENERAL INFORMATION:
? APPLICANT: James A. Wells
? APPLICANT: Brian C. Cunningham
? TITLE OF INVENTION: GROWTH HORMONE VARIANTS
? FILE REFERENCE: 669.12-US-C7
? CURRENT APPLICATION NUMBER: US/10/153,207
? CURRENT FILING DATE: 2002-05-22
? PRIOR APPLICATION NUMBER: 08/479,894
? PRIOR FILING DATE: 1995-06-07
? PRIOR APPLICATION NUMBER: 08/190,723
? PRIOR FILING DATE: 1994-02-02
? PRIOR APPLICATION NUMBER: 07/960,227
? PRIOR FILING DATE: 1992-10-13
? PRIOR APPLICATION NUMBER: 07/875,204
? PRIOR FILING DATE: 1992-04-27
? PRIOR APPLICATION NUMBER: 07/428,066
? PRIOR FILING DATE: 1989-10-26
? PRIOR APPLICATION NUMBER: 07/264,611
? PRIOR FILING DATE: 1988-10-28
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? TYPE: PR
? LENGTH: 191
? ORGANISM: Homo Sapiens
? US-10-153-207-1

Query Match 95.1%; Score 255; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 FETPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 49
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FETPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 48

RESULT 7
US-10-400-377-1
? Sequence 1, Application US/10400377
? Publication No. US20030162949A1
? GENERAL INFORMATION:
? APPLICANT: Cox 171, George N
? APPLICANT: Boldor Biotechnology, Inc.
? TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
? FILE REFERENCE: 4152-1-PJS
? CURRENT APPLICATION NUMBER: US/10/400,377
? CURRENT FILING DATE: 2003-03-26
? PRIOR APPLICATION NUMBER: US/09/462,941
? PRIOR FILING DATE: 2000-01-14
? PRIOR APPLICATION NUMBER: 60/052,516
? PRIOR FILING DATE: 1997-07-14
? NUMBER OF SEQ ID NOS: 41
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 191
? TYPE: PR
? ORGANISM: Homo sapiens
? US-10-400-377-1

Query Match 98.1%; Score 255; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 FETPLSR:FDNMLRAHLRLHQLAFDTYOEFEAYIPKEQKYSFLQNP 49
```

Db 1 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

RESULT 8

US-10-400-708-1
: Sequence 5, Application US/10400729
: Publication No. US20030156855A1
: GENERAL INFORMATION:
: APPLICANT: Cox III, George N
: APPLICANT: Bolder Biotechnology, Inc.
: TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
: FILE REFERENCE: 4152-1-PUS
: CURRENT APPLICATION NUMBER: US/10/400,728
: CURRENT FILING DATE: 2003-03-26
: PRIOR APPLICATION NUMBER: US/09/452,941
: PRIOR FILING DATE: 2000-01-14
: PRIOR APPLICATION NUMBER: 66/052,516
: PRIOR FILING DATE: 1997-07-14
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: Patent in Ver. 2.0
: SEQ ID NO 1
: LENGTH: 191
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-400-708-1

Query Match 98.1%; Score 255; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 1 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 48
|||||

RESULT 9

US-10-153-207-6
: Sequence 6, Application US/10153207
: Publication No. US20030153003A1
: GENERAL INFORMATION:
: APPLICANT: James A. Wells
: APPLICANT: Brian C. Cunningham
: TITLE OF INVENTION: GROWTH HORMONE VARIANTS
: FILE REFERENCE: 569-12-US-07
: CURRENT APPLICATION NUMBER: US/10/153,207
: CURRENT FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: 06/479,864
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 06/190,723
: PRIOR FILING DATE: 1994-02-02
: PRIOR APPLICATION NUMBER: 07/960,227
: PRIOR FILING DATE: 1992-10-13
: PRIOR APPLICATION NUMBER: 07/875,204
: PRIOR FILING DATE: 1992-04-27
: PRIOR APPLICATION NUMBER: 07/428,066
: PRIOR FILING DATE: 1989-10-26
: PRIOR APPLICATION NUMBER: 07/264,611
: PRIOR FILING DATE: 1988-10-28
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: Patent in Ver. 4.0
: SEQ ID NO 6
: LENGTH: 214
: TYPE: PRT
: ORGANISM: Homo Sapiens
US-10-153-207-6

Query Match 98.1%; Score 255; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.5e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

Db 24 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 71
|||||

RESULT 10

US-09-929-918-9
: Sequence 9, Application US/09929918
: Patent No. US20020090678A1
: GENERAL INFORMATION:
: APPLICANT: Kordyum, Vitaliy A.
: APPLICANT: Chernykh, Svitlana I.
: APPLICANT: Slavchenko, Iryna Yu.
: APPLICANT: Vozianov, Olexsandr
: TITLE OF INVENTION: PHAGE-DEPENDENT SUPER PRODUCTION OF
: FILE REFERENCE: PHAGE 006A
: CURRENT APPLICATION NUMBER: US/09/929,918
: CURRENT FILING DATE: 2001-08-15
: PRIOR APPLICATION NUMBER: 09/318,288
: PRIOR FILING DATE: 1999-05-25
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 217
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-929-918-9

Query Match 98.1%; Score 255; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.5e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 27 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74
|||||

RESULT 11

US-09-280-030-66
: Sequence 66, Application US/09280030A
: Patent No. US20010021515A1
: GENERAL INFORMATION:
: APPLICANT: Sato, Seiji
: APPLICANT: Higashikuni, Naoniko
: APPLICANT: Kudo, Toshiyuki
: APPLICANT: Kondo, Masaaki
: TITLE OF INVENTION: DNAs ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
: TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
: FILE REFERENCE: 382-1026
: CURRENT APPLICATION NUMBER: US/09/280,030A
: CURRENT FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: JP10-87339/1998
: EARLIER FILING DATE: 1998-03-31
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: Patent in Ver. 2.0
: SEQ ID NO 66
: LENGTH: 245
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Designated as
: OTHER INFORMATION: an amino acid sequence of MMTsp-MWmp20-TEV-G-GH
US-09-280-030-66

Query Match 98.1%; Score 255; DB 9; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.7e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 55 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 102
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 : Search time 6.57865 seconds
(without alignments)
583.254 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MPTFTLSKLFNMLRAHR.....DEFEKATPKKYSFIQNP 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9618682 residues

Total number of hits satisfying chosen parameters: 283508

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match: 64
Maximum Match: 100
Creating first 45 summaries

Database : PIR76:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	255	98.1	217	1	STBC
2	255	98.1	217	1	STBC
3	228	87.7	217	1	STBC
4	228	87.7	217	1	STBC
5	213	81.9	212	2	STBC
6	213	81.9	212	2	STBC
7	205	78.8	217	2	STBC
8	203	77.3	217	2	STBC
9	197	75.8	215	2	STBC
10	197	75.8	217	1	STBC
11	197	75.8	217	2	STBC
12	161.5	62.1	216	1	STBC
13	160.5	61.3	190	2	STBC
14	159.5	61.3	190	1	STBC
15	159.5	61.3	190	2	STBC
16	159.5	61.3	190	2	STBC
17	159.5	61.3	216	1	STBC
18	159.5	61.3	216	1	STBC
19	159.5	61.3	216	2	STBC
20	159.5	61.3	216	2	STBC
21	159.5	61.3	216	2	STBC
22	159.5	61.3	216	2	STBC
23	156.5	60.2	216	2	STBC
24	155.5	59.8	190	1	STBC
25	150	57.7	216	2	STBC
26	148	56.9	191	2	STBC
27	146	56.2	163	2	STBC
28	144	55.4	190	2	STBC
29	144	55.4	216	2	STBC

30 142.5 54.8 217 1 STBC
31 142.5 54.8 217 1 STBC
32 142.5 54.8 217 1 STBC
33 142.5 54.8 217 2 S32682
34 140 53.8 216 2 S04929
35 132 50.8 190 2 A56816
36 132 50.8 215 2 S11188
37 126 49.2 195 2 S11250
38 128 49.2 215 2 JS0037
39 122 46.9 199 2 B32435
40 116 44.5 183 2 A60623
41 98.5 37.3 87 4 I67761
42 97 37.3 250 2 S11114
43 87 33.5 210 2 S69263
44 87 33.5 210 2 S69262
45 87 33.5 210 2 S02764

ALIGNMENTS

RESULT :

STBC

somatotropin 1 precursor [validated] - human
A: Alternate names: growth hormone 1; hGH-N; pituitary somatotropin
B: Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, s
C: Species: Homo sapiens (man)
D: Date: 24-Apr-1984 #sequence_revision 10-Feb-1995 #text_change 08-Dec-2000
E: Accession: A93731; A32435; A93694; A94247; A90051; A93397; A93778; A91764; A90217
F: Database: PIR76; Moore, D.D.; Goodman, H.M.
G: Nucleic Acids Res. 9, 3719-3730, 1981
A: Title: Human growth hormone DNA sequence and mRNA structure: possible alternative
A: Reference number: A93731; MUID:82014939; PMID:6269091
A: Accession: A93731

A: Molecule type: DNA
A: Residues: 1-217 <DEN>
A: Cross-references: GR:V00520
A: Note: the 20K short form somatotropin lacks residues 58-72 (32-46 in the active h
R: Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Galinas, R.E.; Seebur
Genomics 4, 479-487, 1999
A: Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A: Reference number: A32435; MUID:89307277; PMID:2744760
A: Accession: A32435
A: Molecule type: DNA
A: Residues: 1-217 <CHE>
A: Cross-references: GR:V00520; FIDN:AAA52549.1; PID:g183149
R: Roskam, W.; Rougeon, F.
Nucleic Acids Res. 7, 305-320, 1979
A: Title: Molecular cloning and nucleotide sequence of the human growth hormone stru
A: Reference number: A93694; MUID:80034477; PMID:386281
A: Accession: A93694
A: Molecule type: mRNA
A: Residues: 1-217 <ROS>
A: Cross-references: GR:V00519
A: Note: 35-pro was also found
R: Martini, J.A.; Halliwell, R.A.; Baxter, J.D.; Goodman, H.M.
Science 205, 602-607, 1979
A: Title: Human growth hormone: complementary DNA cloning and expression in bacteria
A: Reference number: A94247; MUID:79203293; PMID:377496
A: Accession: A94247
A: Molecule type: mRNA
A: Residues: 1-217 <MAR>
R: Li, C.H.; Dixon, J.S.; Liu, W.K.
Arch. Biochem. Biophys. 133, 70-91, 1969
A: Title: Human pituitary growth hormone, XIX. The primary structure of the hormone.
A: Reference number: A90648; MUID:69289202; PMID:5810834
A: Contents: annotation
R: Li, C.H.; Dixon, J.S.
Arch. Biochem. Biophys. 146, 233-236, 1971
A: Title: Human pituitary growth hormone, XXII. The primary structure of the hormone
A: Reference number: A90051; MUID:72143935; PMID:5144027
A: Accession: A90051
A: Molecule type: protein

A:Residues: 27-94,96-217 <LIC>
 R:Niall, H.D.
 Nature New Biol. 250, 90-91, 1971.
 A:Title: Revised primary structure for human growth hormone.
 A:Reference number: A93397; NCID:7133765; PMID:5276046
 A:Accession: A93397
 A:Molecule type: protein
 A:Residues: 27-51 <NIA>
 R:Niall, H.D.; Hoan, M.L.; Sayer, R.; Rosenblum, I.Y.; Greenwood, R.C.
 Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971.
 A:Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution.
 A:Reference number: A93778; NCID:71154968; PMID:5276524
 A:Accession: A93778
 A:Molecule type: protein
 A:Residues: 119-120;157-159 <N12>
 R:Niall, H.D.
 in Prolactin and Carcinogenesis, Proc. Fourth Toney's Workshop: Prolactin, Griffiths, R., Ed., Plenum Press, New York, 1977.
 A:Title: The chemistry of the human lactogenic hormones.
 A:Reference number: A94427
 A:Contents: annotation; somatotropin revision
 R:Bewley, I.A.; Dixon, J.S.; Li, C.H.
 Int. J. Pept. Protein Res. 4, 281-287, 1972
 A:Title: Sequence comparison of human pituitary growth hormone, human chorionic somatotropin, and placental lactogen.
 A:Reference number: A91764; NCID:73092628; PMID:4675454
 A:Accession: A91764
 A:Molecule type: protein
 A:Residues: 27-217 <RFW>
 R:Lewis, G.J.; Bonevald, L.F.; Lewis, L.J.
 Biochem. Biophys. Res. Commun. 92, 511-516, 1980
 A:Title: The 20,000-dalton variant of human growth hormone: location of the amino acid sequence.
 A:Reference number: A93227; NCID:60130196; PMID:7336479
 A:Contents: somatotropin, 20K short variant
 A:Accession: A90217
 A:Molecule type: protein
 A:Residues: 46-57;71-80 <RFW>
 R:Chapman, G.E.; Rogers, K.M.; Brittain, T.; Bradshaw, R.A.; Pates, G.J.; Turner, R.; Calkins, J.; Biol. Chem. 256, 2395-2403, 1981
 A:Title: The 20,000 molecular weight variant of human growth hormone. Preparation and sequence.
 A:Reference number: A92311; NCID:6117361; PMID:7462247
 A:Contents: somatotropin, 20K short variant
 A:Accession: A92311
 A:Molecule type: protein
 A:Residues: 27-57;71-79 <CHA>
 R:Singh, R.N.P.; Seavey, B.K.; Lewis, L.J.; Lewis, G.J.
 J. Protein Chem. 2, 425-435, 1983
 A:Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
 A:Reference number: A61466
 A:Accession: A61466
 A:Molecule type: protein
 A:Residues: 27-69 <SIN>
 A:Note: growth hormone 5K peptide has insulin potentiating activity: the physiological form.
 R:Rebbon, V.M.J.; Rao, I.B.; NG, F.
 Biol. Chem. Hoppe-Seyler 371, 423-431, 1990
 A:Title: Identification of the aspartamide structure in a previously reported peptide.
 A:Reference number: S09685; NCID:90334745; PMID:24378673
 A:Accession: S09685
 A:Molecule type: protein
 A:Residues: 27-34, 'L',36-47 <KCB>
 R:De Vos, A.M.; Ullrich, M.; Kossiakoff, A.A.
 Science 255, 306-312, 1992
 A:Title: Human growth hormone and extracellular domain of its receptor: crystal structure.
 A:Reference number: A41728; NCID:52196577; PMID:1254776
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms
 A:Note: the structure of the complex with growth hormone receptor is described.
 R:Gray, G.L.; Balbridge, J.S.; McKewen, K.S.; Heyneker, H.L.; Chang, C.N.
 Gene 39, 247-254, 1985
 A:Title: Periplasmic production of correctly processed human growth hormone in Escherichia coli.
 A:Reference number: I41126; NCID:8637393; PMID:391262
 A:Accession: I41549
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-26 <RES>
 A:Cross-references: GB:M14398; NCID:9183158; PIDN:AA52554.1; PID:q183159

Comment: The gene for this hormone is transcribed only in somatotrophic cells of the pituitary. About 90% of somatotropin is the 22K long form.

Genetics:
 A:Gene: GH3:GH1
 A:Cross-references: GB:M11982; NCIM:119250
 A:Map position: 17q23.1-17q23.3
 A:Insertions: 4/1; 57/3; 97/3; 152/3
 A:Superfamily: prolactin
 A:Keywords: alternative splicing; hormone; pituitary
 F:1-26/Domain: signal sequence status predicted <SIG>
 F:27-217/Product: somatotropin 1, long form #status experimental <SOL>
 F:27-69/Product: growth hormone 5K peptide #status experimental <SKP>
 F:27-57,73-217/Product: somatotropin 1, short form #status experimental <SOS>
 F:79-191,208-215/Disulfide bonds: #status experimental

Query Match 98.1%; Score 255; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2e-24; Mismatches 0; Gaps 0;
 Matches 48; Conservative 0; Indels 0

QY 2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 49
 DB 27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 74

RESULT 2
 167410
 somatotropin - rhesus macaque
 K:Alternate names: growth hormone
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
 C:Accession: I67410; A05094
 R:Goles, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complete
 A:Reference number: I53267; NCID:94008724; PMID:8404617
 A:Accession: I67410
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:I67410; PIDN:AA18842.1; PID:q293115
 R:Li, C.H.; Chung, D.; Lahm, H.W.; Stein, S.
 Arch. Biochem. Biophys. 245, 287-291, 1986
 A:Title: The primary structure of monkey pituitary growth hormone.
 A:Reference number: A05094; NCID:86129460; PMID:3080959
 A:Accession: A05094
 A:Molecule type: protein
 A:Residues: 27-99, 'Q',101-178, 'D',180-217 <LIC>
 A:Note: the monkey species is not identified in the reference
 R:Ruben, M.S.
 Science 125, 883-884, 1957
 A:Title: Preparation of growth hormone from pituitaries of man and monkey.
 A:Reference number: A44774
 A:Contents: annotation; identification of source organism
 A:Superfamily: prolactin

Query Match 98.1%; Score 255; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2e-24; Mismatches 0; Gaps 0;
 Matches 48; Conservative 0; Indels 0

QY 2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 49
 DB 27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 74

RESULT 3
 167410
 somatotropin 2 precursor - human
 K:Alternate names: growth hormone 2; growth hormone variant; hGH-V; placental somatotropin
 C:Species: Homo sapiens (man)
 C:Date: 17-Dec-1982 #sequence_revision 10-Feb-1995 #text_change 21-Jul-2000
 C:Accession: D32435; B28072; A01511; I52104; A60711
 R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seeburg

RESULT 7

167411
 N:Alternat names: rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
 C:Accession: 167411
 R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.B.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomotropin related complementary
 A:Reference number: 153267; MUID:94008724; PMID:8404617
 A:Accession: 167411
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:L16555; NID:q293116; PIDN:AA20480.1; PID:q293117
 C:Superfamily: prolactin

Query Match: 78.8%; Score 205; DB 2; Length 217;
 Best Local Similarity 79.2%; Pred. No. 3,5e-18;
 Matches 38; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 FTPIPLSLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 49
 IIIIIII III I III IIIIIII III IIIIIII III
 DB 27 FTPIPLSLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 74

RESULT 8

167409
 N:Alternat names: rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
 C:Accession: 167409
 R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.B.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomotropin related complementary
 A:Reference number: 153267; MUID:94008724; PMID:8404617
 A:Accession: 167409
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:L16554; NID:q293112; PIDN:AAAL1884.1; PID:h293113
 C:Superfamily: prolactin

Query Match: 77.3%; Score 203; DB 2; Length 217;
 Best Local Similarity 74.5%; Pred. No. 1.1e-17;
 Matches 35; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 3 FTPIPLSLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 49
 IIIIIII III I III IIIIIII III IIIIIII III
 DB 28 PSVPLSKLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 74

RESULT 9

A26449
 N:Alternat names: rhesus macaque
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 28-Jul-1995
 C:Accession: A26449
 R:Hirt, H.; Kneelander, J.; Birnbaum, M.; Chen, E.Y.; Seeburg, P.H.; Eschhardt, N.; Bodo
 DNA 6, 59-70, 1987
 A:Title: The human growth hormone gene locus: structure, evolution, and allelic variation
 A:Reference number: A26449; MUID: 87161235; PMID:3040940
 A:Accession: A26449
 A:Molecule type: DNA
 A:Residues: 1-215 <HR>
 C:Superfamily: prolactin
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-215/Product: chorionamotropin, hCS-3 allele #status predicted <EAF>

Query Match: 75.8%; Score 197; DB 2; Length 215;

Best Local Similarity 80.3%; Pred. No. 3.4e-17;
 Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 TTPVSKLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 48
 IIIIIII III I III IIIIIII III IIIIIII III
 DB 29 TTPVSKLFDNMLRAHRLHQLAFDTYQFEAYIPKEQKYSFLQNP 73

RESULT 10

A26449
 N:Alternat names: chorionic somatomotropin 1; placental lactogen
 C:Species: Homo sapiens (man)
 C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 08-Dec-2000
 C:Accession: C32435; A94422; I52342; A93833; A93192; A90054; A94427; A61283; I55229
 R:Chen, E.Y.; Jaiso, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Geinas, R.E.; Seeburg,
 Genomics 4, 479-497, 1989
 A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
 A:Reference number: A32435; MUID:89307277; PMID:2744760
 A:Accession: C32435
 A:Molecule type: DNA
 A:Residues: 1-217 <CHE>
 A:Cross-references: GB:J03071; NID:gl83148; PIDN:AAAS2551.1; PID:gl83151
 R:Goodman, H.M.; Dekoto, F.; Fiddes, J.C.; Hallowell, R.A.; Page, G.S.; Smith, S.;
 in Mobilization and Reassembly of Genetic Information, Scott, W.A.; Werner, R.; Joshi,
 A:Reference number: A94422
 A:Accession: A94422

A:Molecule type: mRNA
 A:Residues: 1-217 <GOO>
 A:Tanaka, M.; Masuda, N.; Watahiki, M.; Yamakawa, M.; Shimizu, K.; Nagai, J.; Nakas
 Biochem. Int. 16, 287-292, 1988
 A:Title: cDNA cloning of human chorionic somatomotropin-1 mRNA whose transcript
 A:Reference number: I52342; MUID:88209096; PMID:2835050
 A:Accession: I52342
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-3 <TAN>
 A:Cross-references: GB:M35419; NID:g506822
 R:Sherwood, L.M.; Birstein, Y.; Schechter, I.
 Proc. Natl. Acad. Sci. U.S.A. 76, 3819-3823, 1979
 A:Title: Primary structure of the NH-2-terminal extra piece of the precursor to hum
 A:Reference number: A93833; MUID:80034970; PMID:291043
 A:Accession: A93833
 A:Molecule type: protein
 A:Residues: 1-3-26 <SHE>
 A:Experimental source: placenta
 R:Shine, J.; Seeburg, P.H.; Martial, J.A.; Baxter, J.D.; Goodman, H.M.
 Nature 270, 494-499, 1977
 A:Title: Construction and analysis of recombinant DNA for human chorionic somatomam
 A:Reference number: A93192; MUID:78071761; PMID:593366
 A:Accession: A93192

A:Molecule type: DNA
 A:Residues: 50-217 <SHI>
 A:Experimental source: placenta
 R:Li, C.H.; Dixon, J.S.; Chung, D.
 Arch. Biochem. Biophys. 155, 95-110, 1973
 A:Title: Amino acid sequence of human chorionic somatomamotropin.
 A:Reference number: A90054; MUID:73201971; PMID:4712450
 A:Accession: A90054
 A:Molecule type: protein
 A:Residues: 27-217 <SIC>
 A:Experimental source: placenta
 R:Kali, P.D.
 in Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop Prolactin, Griffiths
 A:Title: The chemistry of the human lactogenic hormones.
 A:Reference number: A94427
 A:Accession: A94427
 A:Molecule type: protein
 A:Residues: 27-217 <NIA>
 A:Experimental source: placenta
 R:Nic A Bhaird, N.; Tipton, K.F.
 Biochem. Soc. Trans. 19, 26S, 1991
 A:Title: Catechol-O-methyltransferase from human placenta: purification and some pr

A:Reference number: A61283; MUID:91244006; PMID:2037148
 A:Accession: A61283
 A:Molecule type: protein
 A:Residues: 37-46 <NIG>
 A:Note: chorionamniotropon apparently copurified with placental catechol-O-methyltransferase
 R:Sherwood, L.M.; Handwerger, S.; McLaurin, W.D.; Sander, M.
 Nature New Biol. 233, 59-61, 1972
 A:Title: Amino-acid sequence of human placental lactogen.
 A:Reference number: A93401; MUID:72016313; PMID:5285363
 A:Contents: annotation
 R:Sherwood, L.M.; Handwerger, S.; McLaurin, W.D.; Lunner, M.
 Nature New Biol. 235, 64, 1972
 A:Reference number: A93405
 A:Contents: annotation
 R:Schneider, A.B.; Kowalski, K.; Russell, J.; Sherwood, L.M.
 J. Biol. Chem. 254, 3782-3787, 1979
 A:Title: Identification of the interchain disulfide bonds of dimeric human placental lactogen.
 A:Reference number: A92251; MUID:79173083; PMID:5438159
 A:Contents: annotation; dimeric disulfide bonds
 R:Selby, M.J.; Partridge, A.; Baxter, J.D.; Bell, G.F.; Eberhardt, N.L.
 J. Biol. Chem. 259, 13131-13138, 1984
 A:Title: Analysis of a major human chorionic somatomammotropin gene. Evidence for two transcripts.
 A:Reference number: 155229; MUID:85030425; PMID:6206192
 A:Accession: 155229
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GR:K03401; NID:9181120; PIDN:AAA52155.1; PID:q161121
 R:Seeburg, P.H.; Shine, J.; Harada, J.A.; Gillich, A.; Goodman, H.
 Trans. Assoc. Am. Physicians 90, 109-116, 1977
 A:Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.
 A:Reference number: 159658; MUID:78-60787; PMID:611657
 A:Accession: 159658
 A:Status: translated from GR/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 160-217 <R22>
 A:Cross-references: GB:M26118; NID:9181124; PIDN:AAA5521.1; PID:q161125
 C:Genetics:
 A:Gene: GRN/CSH
 A:Cross-references: GR:113084; OMIM:150200
 A:Map position: 17q22-17q24
 A:Introns: 4/1; 57/3; 97/3; 152/4
 C:Superfamily: prolactin
 C:Keywords: Hormone; Placenta
 F:1-26/Domain: signal sequence #status experimental <SIG>
 F:27-217/Product: chorionamniotropon A #status experimental <MAT>
 F:79-191/Disulfide bonds: #status experimental
 F:208-215/Disulfide bonds: (in monomeric form) #status experimental
 F:208/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental
 F:215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental
 Query Match 75.8%; Score 197; DB 1; Length 217;
 Best Local Similarity 80.0%; Pred. No. 3.5e-17;
 Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 4 TPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 48
 DB 29 TVPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLHD 73
 RESULT 11
 E32435
 chorionamniotropon B precursor - human
 N:Alternate names: chorionic somatomammotropin 2
 C:Species: Homo sapiens (man)
 C:Date: 29-Dec-1989 #sequence_revision 29-Dec-1989 text_change 16-Jul-1999
 C:Accession: E32435
 R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Herrera-Saidana, H.A.; Gelinas, A.E.; Seeburg, P.
 Genomics 4, 479-497, 1989
 A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.
 A:Reference number: A32435; MUID:89307277; PMID:2744760
 A:Accession: E32435
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-217 <CHE>
 A:Cross-references: GB:J03071; NID:9183148; PIDN:AAA52553.1; PID:q183153
 C:Genetics:
 A:Gene: GDF/CSH2
 A:Cross-references: GR:1119813; OMIM:118820
 A:Map position: 17q22-17q24
 C:Superfamily: prolactin
 Query Match 75.8%; Score 197; DB 2; Length 217;
 Best Local Similarity 80.0%; Pred. No. 3.5e-17;
 Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 4 TPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 48
 DB 29 TVPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLHD 73
 RESULT 12
 STMS
 somatotropon precursor - mouse
 N:Alternate names: growth hormone
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 text_change 28-May-1999
 C:Accession: E23911
 R:Linnar, D.L.H.; Talamantes, F.
 J. Biol. Chem. 260, 9574-9579, 1985
 A:Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression.
 A:Reference number: A92548; MUID:85261358; PMID:2991252
 A:Accession: E23911
 A:Molecule type: mRNA
 A:Residues: 1-216 <LIN>
 A:Cross-references: GB:X02891; GB:K03232; NID:951067; PIDN:CAA26650.1; PID:q51068
 C:Superfamily: prolactin
 C:Keywords: anterior pituitary; growth factor; hormone
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-216/Product: somatotropon #status predicted <SIG>
 F:74-189,206-214/Disulfide bonds: #status predicted
 Query Match 62.2%; Score 161.5; DB 1; Length 216;
 Best Local Similarity 68.1%; Pred. No. 9.4e-13;
 Matches 32; Conservative 5; Mismatches 8; Indels 1; Gaps 1;
 QY 2 FTTPILSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 48
 DB 27 FTAMPILSLFSDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 72
 RESULT 13
 PN0140
 somatotropon - sei whale
 N:Alternate names: growth hormone
 C:Species: Balaeocetora borealis (sei whale)
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 text_change 07-May-1999
 C:Accession: PN0140
 R:Yudaev, N.A.; Pankov, Y.A.; Bulatov, A.A.; Osipova, T.A.
 Biochimia 47, 1059-1069, 1982
 A:Title: Amino acid sequence of sei whale somatotropon.
 A:Reference number: PN0140; MUID:83000569; PMID:7115813
 A:Accession: PN0140
 A:Molecule type: protein
 A:Residues: 1-190 <YUD>
 A:Note: article in Russian with English abstract
 C:Superfamily: prolactin
 C:Keywords: growth factor; hormone
 F:52-163,180-188/Disulfide bonds: #status predicted
 Query Match 61.7%; Score 160.5; DB 2; Length 190;
 Best Local Similarity 68.1%; Pred. No. 1.1e-12;
 Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
 QY 2 FTTPILSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 48
 DB 27 FTAMPILSLFSDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQN 72

DB 1 FPMPLSLFANAVLRAGHRLHQLAFDTYQSEFEAYTPEGRYS-ION 46

RESULT 14

STHO

Somatotropin - horse
N:Alternate names: growth hormone
C:Species: Equus caballus (domestic horse)
C>Date: 13-Jul-1981 sequence revision 13-Jul-1981 text_change 24-Aug-1996
C:Accession: A91395; A91393; A90240; A01534
R:Zakin, M.M.; Poskus, E.; Langston, A.A.; Ferrara, P.; Santome, J.A.; Dellacha, J.A.;
Int. J. Pept. Protein Res. 8, 435-444, 1976
A:Title: Primary structure of equine growth hormone.
A:Reference number: A91395; MUID:7700541C; PMID:965151
A:Accession: A91395
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakin, M.M.; Poskus, E.; Dellacha, J.A.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 34, 353-355, 1973
A:Title: The amino acid sequence of equine growth hormone.
A:Reference number: A91395; MUID:74020362; PMID:4747549
A:Accession: A91395
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakin, M.M.; Poskus, E.; Dellacha, J.A.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 25, 77-82, 1972
A:Title: Amino acid sequences around the cystine residues in equine growth hormone.
A:Reference number: A91383
A:Accession: A91383
A:Molecule type: protein
A:Residues: 42-59; 157-190 <ZAK>
R:Oliver, L.; Hartree, A.S.
Biochem. J. 109, 19-24, 1968
A:Title: Amino acid sequences around the cystine residues in horse growth hormone.
A:Reference number: A90240; MUID:68383390; PMID:4976109
A:Accession: A90240
A:Molecule type: protein
A:Residues: 176-190 <GLI>
C:Superfamily: prolactin
C:Keywords: hormone; pituitary
P:52-163,180-168/disulfide bonds: *status experimental

Query Match 61.3%; Score 159.5; DB 2; Length 190;
Best Local Similarity 65.1%; Pred. No. 1.4e-12;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 PPTPLSLFDNMLRAHRLHQLAFDTYQSEFEAYTPEGRYS-ION 46

DB 1 FPMPLSLFANAVLRAGHRLHQLAFDTYQSEFEAYTPEGRYS-ION 46

RESULT 15

JK0219

Somatotropin - African elephant
N:Alternate names: growth hormone
C:Species: Loxodonta africana (African elephant)
C>Date: 03-Aug-1992 sequence revision 03-Aug-1992 text_change 15-Nov-1996
C:Accession: JK0219
R:Hulmes, J.D.; Miedel, M.C.; Li, C.H.; Fan, Y.C.E.;
Int. J. Pept. Protein Res. 33, 356-372, 1989
A:Title: Primary structure of elephant growth hormone.
A:Reference number: JK0219
A:Accession: JK0219
A:Molecule type: protein
A:Residues: 1-190 <HUL>
A:Experimental source: pituitary gland
C:Superfamily: prolactin
P:1-190/Product: somatotropin *status experimental <MA>

Query Match 61.3%; Score 159.5; DB 2; Length 190;
Best Local Similarity 65.1%; Pred. No. 1.4e-12;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 PPTPLSLFDNMLRAHRLHQLAFDTYQSEFEAYTPEGRYS-ION 48
DB 1 FPMPLSLFANAVLRAGHRLHQLAFDTYQSEFEAYTPEGRYS-ION 46

Search completed: September 15, 2003, 12:04:19
Job time : 9.07885 secs

GenCore version 5.1.1
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:00 : Search time 4.56631 Seconds
(without alignments)
504,635 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MFPTPLSRKFDNMKRNK.....QRPPEAYIPKSKYKFFQNP 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match (%)
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	255	98.1	217	1	SOMA_HUMAN
2	255	98.1	217	1	SOMA_MACMU
3	255	98.1	217	1	SOMA_PANTR
4	249	95.8	217	1	SOMA_CALJA
5	249	95.8	217	1	SOMA_SABR
6	236	90.8	217	1	SCM2_PANTR
7	228	87.7	217	1	SCM2_HUMAN
8	199	76.5	217	1	SOMA2_MACMU
9	197	75.8	217	1	PILL_HUMAN
10	161.5	62.1	216	1	SOMA_MOUSE
11	160.5	61.7	190	1	SOMA_BALRO
12	159.5	61.3	190	1	SOMA_LOXAF
13	159.5	61.3	190	1	SOMA_VULVU
14	159.5	61.3	216	1	SOMA_CANPA
15	159.5	61.3	216	1	SOMA_FELCA
16	159.5	61.3	216	1	SOMA_HORSE
17	159.5	61.3	216	1	SOMA_MESAU
18	159.5	61.3	216	1	SOMA_PIG
19	159.5	61.3	216	1	SOMA_RABIT
20	159.5	61.3	216	1	SOMA_RAT
21	159.5	61.3	217	1	SOMA_GAUSE
22	159.5	61.3	217	1	SOMA_NYCPY
23	159.5	61.3	216	1	SOMA_MUSVI
24	155.5	59.8	190	1	SOMA_IAMPA
25	150	57.7	216	1	SOMA_YELCA
26	148	56.9	191	1	SOMA_CHEMF
27	145	55.8	215	1	SOMA_MONDO
28	145	55.8	215	1	SOMA_TRIVU
29	144	55.4	190	1	SOMA_ACIQU
30	144	55.4	190	1	SOMA2_ACIQU
31	144	55.4	216	1	SOMA_CHICK
32	142.5	54.8	217	1	SOMA_POVIN
33	142.5	54.8	217	1	SOMA_CERFEL

14	142.5	54.8	217	1	SOMA_SHEEP
35	142	54.6	217	1	SOMA_STRCA
36	140	53.8	190	1	SOMA_CRONO
37	140	53.8	216	1	SOMA_ANAPL
38	135.5	52.1	217	1	SOMA_BUBBU
39	132	50.8	215	1	SOMA_RANCA
40	125	48.1	211	1	SOMA_LEPOS
41	122	46.9	214	1	SOMA_XENLA
42	116	44.6	183	1	SOMA_XENLA
43	112	43.1	206	1	SOMA_PROAN
44	111	42.7	213	1	SOMA_BUFMA
45	104	49.0	208	1	SOMA_XENLA

RESULT 1
SOMA_HUMAN

AC SOMA_HUMAN STANDARD: PRT: 217 AA.
PE1241: Q14405; Q16531; Q9HB22; Q90K57; Q90NL5;
21-JUL-1986 (rel. 01, Created)
01-1-MAR-1992 (rel. 21, Last sequence update)
01-15-SEP-2003 (rel. 42, Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone :).
GN GH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=80034477; PubMed=39628;
RA Roskam W., Rougeon F.;
RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene".
RI Nucleic Acids Res. 7:305-320(1979).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=79203253; PubMed=377456;
RA Martini J.A., Ballewell R.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression in bacteria".
RI Science 205:632-637(1979).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.
RX MEDLINE=82014939; PubMed=6269091;
RA Denoto F.M., Moore D.D., Goodman H.M.;
RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing".
RI Nucleic Acids Res. 9:3719-3730(1981).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=84182510; PubMed=7169009;
RA Seeburg P.H.;
RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone".
RI DNA 1:239-249(1982).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=89307277; PubMed=2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
RT "The human growth hormone locus: nucleotide sequence, biology, and evolution".
RI Genomics 4:479-497(1989).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX TISSUE=Pituitary;
RA Gu J., Huang O.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;
RT "A novel gene expressed in human pituitary".
RI Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RA SEQUENCE FROM N.A. (ISOFORM 4).
 RP TISSUE-Pituitary; PubMed:1334946;
 RX MEDLINE-20402571; PubMed:1334946;
 RA Gu R.-M., Han Z.-G., Song H.-C., Peng Y.-D., Huang G.-H., Ren G.-X.,
 RA Hu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-B.,
 RA Gu S.-W., Dai M., Mao Y.-P., Gao G.-F., Kong B., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhao G.-K., Wu T.-Y.,
 RA Huang G.-Y., Chen Z., Chen K.-D., Chen G.-L.,
 RT "Gene expression profiling in the human hypothalamus pituitary-adrenal
 RT axis and full-length cDNA cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:9543-9548(2000).
 RN [18]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE-86137393; PubMed:39 2251;
 RA Gray G.L., Badierde J.S., McKewen K.S., Heyneker H.L., Chuaq C.N.,
 RT "Periplasmic production of correctly processed human growth hormone in
 RT *Escherichia coli*: natural and bacterial signal sequences are
 RT interchangeable.";
 RL Gene 39:247-254(1985).
 RN [19]
 RP SEQUENCE OF 27-217.
 RX MEDLINE-69289202; PubMed:58 0334;
 RA Li C.H., Dixon J.S., Liu W.K.,
 RT "Human pituitary growth hormone. XIX. The primary structure of the
 RT hormone.";
 RL Arch. Biochem. Biophys. 133:70-91(1969).
 RN [16]
 RP SEQUENCE OF 27-217 AND REVISIONS.
 RX MEDLINE-72143935; PubMed:544027;
 RA Li C.H., Dixon J.S.,
 RT "Human pituitary growth hormone. 32. The primary structure of the
 RT hormone: revision.";
 RL Arch. Biochem. Biophys. 146:233-246(1971).
 RN [11]
 RP REVISION
 RX MEDLINE-73092028; PubMed:4675454;
 RA Bewley T.A., Dixon J.S., Li C.H.,
 RT "Sequence comparison of human pituitary growth hormone, human
 RT chorionic somatomotropin, and ovine pituitary growth and
 RT lactogenic hormones.";
 RL Int. J. Pept. Protein Res. 4:281-287(1972).
 RN [12]
 RP SEQUENCE OF 27-61 AND 102-134.
 RX MEDLINE-7139745; PubMed:5279045;
 RA Niall H.D.,
 RT "Revised primary structure of human growth hormone.";
 RL Nature New Biol. 230:96-97(1971).
 RN [13]
 RP REVISIONS TO 119-120 AND 157-159.
 RX MEDLINE-71153968; PubMed:5279528;
 RA Niall H.D., Hodan M.E., Sauer R., Rosenblum I.Y., Greenwood F.C.,
 RT "Sequences of pituitary and placental lactogenic and growth hormones:
 RT evolution from a primordial peptide by gene reduplication.";
 RL Proc. Natl. Acad. Sci. U.S.A. 68:865-869(1971).
 RN [14]
 RP REVISION.
 RX MEDLINE-71153968; PubMed:5279528;
 RA Niall H.D.,
 RT "The chemistry of the human lactogenic hormones.";
 RL (In) Griffiths K. (eds.),
 RL Prolactin and carcinogenesis, Proc. Fourth Taceous Workshop Prolactin
 RL pp.13-20, Alpha Omega Alpha Press, Cardiff (1972).
 RN [15]
 RP SEQUENCE OF 27-79 (ISOFORM 2).
 RX MEDLINE-8117361; PubMed:7462247;
 RA Chapman G.E., Rogers K.M., Brittain J., Bradshaw E.A., Paton G.J.,
 RA Turner C., Gary P.D., Crane Robinson G.,
 RT "The 20,000 molecular weight variant of human growth hormone:
 RT Preparation and some physical and chemical properties.";
 RL J. Biol. Chem. 256:2395-2401(1981).
 RN [16]
 RP SEQUENCE OF 46-80 (ISOFORM 2).
 RX MEDLINE-80130196; PubMed:7356479;
 RA Lewis U.G., Borewald L.F., Lewis L.J.,
 RT "The 20,000-dalton variant of human growth hormone: location of the
 RT amino acid deletions."; Commun. 92:511-516(1980).
 RN [17]
 RP REAMIFICATION OF G1N-153 AND ASN-178.
 RX MEDLINE-82052997; PubMed:7028740;
 RA Lewis U.G., Stapp R.N., Borewald L.F., Seavey B.K.,
 RT "Altered proteolytic cleavage of human growth hormone as a result of
 RT reamidation."; J. Biol. Chem. 256:11645-11650(1981).
 RN [18]
 RP REVIEW.
 RX MEDLINE-99321812; PubMed:10393484;
 RA Bauwenn G.,
 RT "Growth hormone heterogeneity in human pituitary and plasma.";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RN [19]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-8819073; PubMed:3447173;
 RA Cohen F.R., Kuntz I.D.,
 RT "Prediction of the three-dimensional structure of human growth
 RT hormone.";
 RL Proteins 2:162-166(1987).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE-92196577; PubMed:1549776;
 RA de Vos A.M., Utsch M., Kossiakoff A.A.,
 RT "Human growth hormone and extracellular domain of its receptor:
 RT crystal structure of the complex.";
 RL Science 255:306-312(1992).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE-95075462; PubMed:7984244;
 RA Somers W., Utsch M., de Vos A.M., Kossiakoff A.A.,
 RT "The X-ray structure of a growth hormone-prolactin receptor complex.";
 RL Nature 372:478-481(1994).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE-97113023; PubMed:8943276;
 RA Pavlovsk A.G., Wlodawer A.,
 RT "The crystal-structure of wild-type growth-hormone at 2.5-A
 RT resolution.";
 RL Protein Pept. Lett. 2:333-340(1995).
 RN [23]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE-97113023; PubMed:8943276;
 RA Sundstrom M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,
 RT "Crystal structure of an antagonist mutant of human growth hormone,
 RT G120R, in complex with its receptor at 2.9-A resolution.";
 RL J. Biol. Chem. 271:32197-32203(1996).
 RN [24]
 RP VARIANT CYS-105.
 RX MEDLINE-99348093; PubMed:10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.C.,
 RA Lander E.S.,
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [25]
 RP ERRATUM.
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.C.,
 RA Lander E.S.,
 RL Nat. Genet. 23:373-373(1999).
 RN [26]
 RP FUNCTION: Plays an important role in growth control. Its major
 RP role in stimulating body growth is to stimulate the liver and
 RP other tissues to secrete IGF-1. It stimulates both the
 RP differentiation and proliferation of myoblasts. It also stimulates

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatotropin precursor (Growth hormone).
 GN GH1.
 OS Callithrix jacchus (Common marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callithricidae;
 OC Callithrix.
 OX NCBI_TaxID=9483;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis O.C., Wallis M.;
 RT "Cloning and characterisation of a putative growth hormone encoding
 RL gene from the marmoset (Callithrix jacchus)."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth in growth control. It also stimulates
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AJ297563; CAC3481.1;
 CC HSP: P01241; I222.
 CC InterPro: IPR001400; Somatotropin.
 CC Pfam: PF00103; hormone; 1.
 CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
 CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
 CC Hormone; Pituitary; Signal.
 CC SIGNAL 1 26 BY SIMILARITY.
 CC CHAIN 27 217 SOMATOTROPIN.
 CC DISULFID 79 191 BY SIMILARITY.
 CC DISULFID 208 215 BY SIMILARITY.
 CC SEQUENCE 217 AA; 24959 MW; E:02151A12CE56:92 CRC64;
 CC
 CC Query Match 95.8%; Score 249; DB 1; Length 217;
 CC Best local Similarity 97.9%; Pred. No. 3.7e-24;
 CC Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 2 FPTPLSLRLENAMLRARHLRLHQLAFDTYCEFEAYIPKEOKYSFLONP 49
 CC 27 FPTPLSLRLENAMLRARHLRLHQLAFDTYCEFEAYIPKEOKYSFLONP 74
 CC
 CC RESULT 5
 CC SOMA_SAIIB STANDARD; PRT; 217 AA.
 CC AC P58433.
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Somatotropin precursor (Growth hormone).
 GN GH1.
 OS Saimiri boliviensis boliviensis (Bolivian squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=39432;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21265430; PubMed=11371582;
 RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
 RT "Episodic evolution of growth hormone in primates and emergence of the
 RT species specificity of human growth hormone receptor."
 RL Mol. Biol. Evol. 18:945-953(2001).

CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth in growth control. It also stimulates
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF339560; AAK52287.1;
 CC InterPro: IPR001400; Somatotropin.
 CC Pfam: PF00103; hormone; 1.
 CC PROSITE: PS00336; SOMATOTROPIN.
 CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
 CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
 CC Hormone; Pituitary; Signal.
 CC SIGNAL 1 26 BY SIMILARITY.
 CC CHAIN 27 217 SOMATOTROPIN.
 CC DISULFID 79 191 BY SIMILARITY.
 CC DISULFID 208 215 BY SIMILARITY.
 CC SEQUENCE 217 AA; 24864 MW; 9515289992C529F7 CRC64;
 CC
 CC Query Match 95.8%; Score 249; DB 1; Length 217;
 CC Best local Similarity 97.9%; Pred. No. 3.7e-24;
 CC Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 2 FPTPLSLRLENAMLRARHLRLHQLAFDTYCEFEAYIPKEOKYSFLONP 49
 CC 27 FPTPLSLRLENAMLRARHLRLHQLAFDTYCEFEAYIPKEOKYSFLONP 74
 CC
 CC RESULT 6
 CC SOM2_PANIR STANDARD; PRT; 217 AA.
 CC AC P38757;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
 CC hormone) (Growth hormone 2).
 GN GH2.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 OX NCBI_TaxID=9596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Revol A., Esquivel D., Santiago D., Barrera-Saidana H.;
 RT "Independent duplication of the growth hormone gene in three
 RT Anthropoid lineages."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth in growth control. It stimulates both the
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed in the placenta.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to licensed@isb-sib.ch).

DR EMBL: AF374233; AAL72285.1.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Placenta; Signal; Glycoprotein.
 FT SIGNAL; 1-26 BY SIMILARITY.
 FT CHAIN; 27-217 GROWTH HORMONE VARIANT.
 FT DISULFID; 79-191 BY SIMILARITY.
 FT DISULFID; 208-215 BY SIMILARITY.
 SQ SEQUENCE 217 AA; 24990 MW; 1592A42607577EE CR004;

Query Match 90.8%; Score 236; D5 D; Length 217;
 Best Local Similarity 93.8%; Prod. No. 1.6e-22;
 Matches 45; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 FPTIPSLRFLNAMLRLHQLAFTYQFEFAYIFKECKYSFLQNP 49
 DB 27 FPTIPSLRFLNAMLRLHQLAFTYQFEFAYIFKECKYSFLQNP 74

RESULT 7

ID SOM2_HUMAN STANDARD; PRT: 217 AA.
 AC P01242; P09587;
 DT 21-JUL-1986 (Rel. 01. Created);
 DI 28-FEB-2003 (Rel. 42. Last sequence update);
 DI 15-SEP-2003 (Rel. 42. Last annotation update);
 DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
 DE hormone) (Growth hormone 2).
 GN GH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RX MEDLINE=83182010; PubMed=7159509;
 RA Seeburg P.H.;
 RT "The human growth hormone gene family: nucleotide sequences show
 RT recent divergence and predict a new polypeptide hormone.";
 RL DNA 1:239-249(1982).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=88243759; PubMed=3179057;
 RA Cooke N.E., Ray J.J., Emery J.G., Liebhafner S.A.;
 RT "Two distinct species of human growth hormone-variant mRNA in the
 RT human placenta predict the expression of novel growth hormone
 RT proteins";
 RL J. Biol. Chem. 263:9001-9006(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=89024984; PubMed=2450050;
 RA Igout A., Scippo M.L., Frankenne P., Hennein G.;
 RT "Cloning and nucleotide sequence of placental hGH-V cDNA.";
 RL Arch. Int. Physiol. Biochim. 96:63-67(1988).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89307277; PubMed=2744760;
 RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
 RA Gellinas R.E., Seeburg P.H.;
 RT "The human growth hormone locus: nucleotide sequence, biology, and
 RT evolution.";
 RL Genomics 4:479-497(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12177932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,

KA Katscher R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 PA Allschul S.F., Zeeberg B., Bietow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant I.L., Scheetz T.E.,
 PA Brownstein M.J., Udgin T.B., Yoshiyuki S., Carninci P., Prange C.,
 KA Kana S.S., Loguc-laco N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Rosak S.A., McWayne P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wealey K.C., Hale S., Garcia A.M., Gay L.J., Hu-Yk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gramwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalusz D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 PR Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP REV:EW.
 MEDLINE=95321812; PubMed=10393484;
 BA Baumann G.;
 RT "Growth hormone heterogeneity in human pituitary and plasma.";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 CC -- FUNCTION: Plays an important role in growth control. It's major
 CC role in stimulating body growth is to stimulate the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues.
 CC -- SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer. disulfide-
 CC linked or non-covalently associated, in homopolymeric and
 CC heteropolymeric combinations. Can also form a complex either with
 CC GHBP or with the alpha2-macroglobulin complex.
 CC -- SUBCELLULAR LOCATION: Secreted.
 CC -- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=GH-V1;
 CC "IsoId=P01242-1; Sequence=Displayed;
 CC Name=2; Synonyms=GH-V2;
 CC "IsoId=P01242-2; Sequence=VSP_006203;
 CC Note=No experimental confirmation available;
 CC -- TISSUE SPECIFICITY: Expressed in the placenta.
 CC -- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to licensed@isb-sib.ch).

CC EMBL: K00470; AAA98619.1;
 DR EMBL: J03756; AA859547.1;
 DR EMBL: J03756; AA859548.1;
 DR EMBL: M38451; AAA35891.1;
 DR EMBL: J03071; AAA52352.1;
 DR EMBL: BCC20760; AAH20760.1;
 DR PIR: A28072; SIHV2.
 DR PIR: D32435; SIHVJ.
 DR HSSP: P01241; 1A22.
 DR Genew: HGNC:4262; GH2.
 DR MIM: 139240;
 DR GO: GO:0005180; Peptide hormone; IAS.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Placenta; Signal; Glycoprotein; Alternative splicing;
 KW Polymorphism. 1
 FT SIGNAL.


```

EMBL: J00118; AAA98621.1; ..
DR EMBL: BC002717; AAA02717.1; ..
DR EMBL: BC005921; AAA05921.1; ..
DR EMBL: BC020756; AAA20756.1; ..
DR FJR: A26449; A26449.
DR FJR: C32435; LC80C.
DR HSSP: POL241; IA22.
DR Genew: HGNC:244C; CSH1.
DR Genew: HGNC:2441; CSH2.
DR MIM: 150200; ..
DR GO: GO:0007565; P:pregnancy; TAS.
DR InterPro: IPR01400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PROSITE: PS00366; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Placenta; Multigene family; Signal.
FT STGNAL; 1 26
FT CHAIN; 27 217
FT DISULFID; 75 191
FT DISULFID; 208 215
FT DISULFID; 208 208
FT DISULFID; 215 215
FT VARIANT; 3 3
FT VAR-ANT; 134 105
FT CONFLICT; 84 84
FT CONFLICT; 95 95
FT CONFLICT; 116 116
FT CONFLICT; 134 136
FT CONFLICT; 134 136
FT SEQUENCE; 217 AA; 25320 MW; 235B0DC7A713F431 CRC64;
FT VAR-ANT; 134 105
FT CONFLICT; 84 84
FT CONFLICT; 95 95
FT CONFLICT; 116 116
FT CONFLICT; 134 136
FT CONFLICT; 134 136
FT SEQUENCE; 217 AA; 25320 MW; 235B0DC7A713F431 CRC64;
Query Match 75.8%; Score 197; DB 1; Length 217;
Best Local Similarity 90.0%; Pred. No. 1.3e-17;
Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
CO 4 TPPLSELDNNAKLRAHRLHQAFTYQYEFEEATIPKQKYSPLON 48
ID 29 TPPLSELDNNAKLRAHRLHQAFTYQYEFEEATIPKQKYSPLON 73
RESULT 10
SCHEMA_MOUSE SCHEMA_MOUSE
AC SCHEMA_MOUSE STANFORD; FR7; 216 AA.
AC P06880;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Fast sequence update)
DT 23-FEB-2003 (Rel. 41, Last annotation update)
DR Somatotropin precursor (Growth hormone).
DR GH1 CR GH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC NCBI_TaxID:10050;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE-85261358; PubMed-2391252;
RA Linder D.I.H., Idanmantes F.;
RT "Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
expression of these mRNAs during pregnancy.";
RL J. Biol. Chem. 260:9574-9579(1985).
RN 12;
RP SEQUENCE FROM N.A.
RC S-RAIN-F2TDU; TISSUE: Liver;
RX MEDLINE-96194603; PubMed-8647448;
RA Das P., Meyer L., Seyfert H.-M., Brockman G., Schwerin M.;
RT "Structure of the growth hormone-encoding gene and its promoter in
mice.";
RL Gene 169:209-213(1996).
CC -1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC

```

```

LR HSSP; P01247; IAX::
IP InterPro: IPR001400; Somatotropin.
PF Pfam: PF00103; Hormone_1.
PR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; 163 BY SIMILARITY.
FT DISULFID 52 163 BY SIMILARITY.
FT DISULFID 180 188 BY SIMILARITY.
SQ SEQUENCE 190 AA; 21635 MW; 09BF6D814A75D6 CRC64;

Query Match 61.7%; Score 150.5; DB 1; Length 190;
Rest Local Similarity 66.1%; Pred. NO. 5-5e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps

QY 2 FTPIPLSRIFENAMRAHRLHQIADPTVOEERAYITPEKOKYSFLQN 48
   II IIII :IIIII :II :IIIIII :IIII :IIII
IQ 1 FPAEMPLSSIFANAVLIRAGHLIELAADTYKEFERAYIEGGORY-FLQN 46

RESULT 12
SOMA_VULVU STANDARD; PRT: 190 AA.
AC AC 330362;
DT 01-FEB-1991 (Rel. 17, Created)
DI 01-FEB-1991 (Rel. 17, Last sequence update)
DT 24-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin (Growth hormone).
GN Ghil.
OS Loxodonta africana (African elephant).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Loxodonta.
CX NCBI_TaxID=9785;
RK RK 1;;
RP SEQUENCE.
RA Hulmes J.D., Miedel M.C., Li C.H., Pan Y.C.E.:
RL Int. J. Pept. Protein Res. 33:368-372(1989).
CC CC -I- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CR CR PIR: JK0219; JK0219.
CR HSSP; P01246; IBST.
GR InterPro: IPR001400; Somatotropin.
PF Pfam: PF00103; hormone_1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
AW Hormone; Pituitary.
FT DISULFID 52 163 BY SIMILARITY.
FT DISULFID 180 188 BY SIMILARITY.
SQ SEQUENCE 190 AA; 21761 MW; 05B860313D8741F2 CRC64;

Query Match 61.3%; Score 159.5; DB 1; Length 190;
Rest Local Similarity 66.1%; Pred. NO. 5-5e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps

QY 2 FTPIPLSRIFENAMRAHRLHQIADPTVOEERAYITPEKOKYSFLQN 48
   II IIII :IIIII :IIII :IIIIII :IIII :IIII
QB 1 FPAEMPLSSIFANAVLIRAGHLIELAADTYKEFERAYIEGGORY-IQN 46

RESULT 13
SOMA_VULVU STANDARD; PRT: 190 AA.
AC AC 310756;
DT 01-JUL-1989 (Rel. 11, Created)
DI 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin (Growth hormone).
```


CC -1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, European
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see <http://www.1st-sib.ch/permissions/>
CC or send an email to license@sib-sb.ch).
CC
CC EMBL: G25973; AAA57294.1; ..
CC EMBL: U13390; AAA96142.1; ..
CC PIR: JC4632; JC4632.
CC HSP: P01246; IBSI.
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF06103; Hormone; 1.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
FT CONFLICT 7 7 N -> T (IN REF. 2).
FT CONFLICT 26 26 T -> A (IN REF. 2).
FT CONFLICT 159 159 G -> A (IN REF. 2).
FT CONFLICT 181 181 L -> P (IN REF. 2).
SQ SEQUENCE 216 AA; 24454 MW; 65820239A7D292C6 CRC64;

Query Match 61.3%; Score 159.5; DB 1; length 216;
Best Local Similarity 68.1%; Pred. No. 6.3e-13;
Matches 32; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Cy 2 FPIIPLSRLFDNAMLRAHLHQLAFTYCEFEERAYTRKSKYFLQN 48
1 : ||| : |||| : || : ||||| : || : ||| :
Db 27 FPMPLSLFANAVIRACHLHQLAATYKEPERAYIDEGGRYS-IGN 72

Search completed: September 15, 2003, 12:01:29
Job time : 5.56631 secs

GenCore version 3.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:32 : Search time 15.1434 seconds
(without alignments)
660.520 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MFPTPLSLRFDNMLRAHR.....QEFEEAYTPKQKYSFLQNP 49

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 640525 seqs, 253652604 residues

Total number of hits satisfying chosen parameters: 410525

Minimum DB seq length: 6

Maximum DB seq length: 200000000

Post-processing: Minimum Match 10

Maximum Match 1000

Listing first 45 summaries

Database :

- 1: SP_ARCHAEA*
- 2: SP_BACTERIA*
- 3: SP_FUNGI*
- 4: SP_HUMAN*
- 5: SP_INVERTEBRATE*
- 6: SP_MAMMAL*
- 7: SP_MITE*
- 8: SP_ORGANELLE*
- 9: SP_PLAGE*
- 10: SP_PLANT*
- 11: SP_PROTECT*
- 12: SP_VIRUS*
- 13: SP_VERTEBRATE*
- 14: SP_UNCLASSIFIED*
- 15: SP_VIRUS*
- 16: SP_BACTERIA*
- 17: SP_ARCHAEA*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	95.8	217	6	Q6WNEQ
2	228	87.7	245	4	Q14644
3	213	81.9	212	6	Q07368
4	213	81.9	217	6	Q07367
5	202	77.3	217	6	Q07369
6	197	75.8	217	4	Q14407
7	195	75.0	217	6	Q8WND9
8	170	65.4	217	6	Q8M174
9	160.5	61.7	216	11	Q70615
10	159.5	61.3	52	6	Q5TV91
11	159.5	61.3	216	6	Q8M173
12	159.5	61.3	216	6	Q8HYE5
13	155.5	59.8	216	11	Q9R2C3
14	154	59.2	216	11	Q5JRM4
15	152	58.5	178	6	Q65M75
16	149.5	57.5	204	6	Q95205

17	147	56.5	202	4	Q14643
18	146	56.2	178	6	Q95M76
19	144	55.4	190	11	Q9JKG0
20	143.5	55.2	192	6	Q9TU21
21	142.5	54.8	192	6	Q9TOW9
22	141	54.2	217	6	Q8M175
23	138.5	53.3	217	6	Q28957
24	138.5	53.3	217	6	Q9BEC0
25	138.5	53.3	217	6	Q9BEB9
26	137	52.7	40	6	Q9TR19
27	132	50.8	218	13	Q9PU72
28	128	49.2	195	13	Q91386
29	123.5	47.5	143	6	Q95240
30	122	46.9	149	4	Q14406
31	119	45.8	63	13	Q8QG85
32	107	41.2	53	6	Q19034
33	104	40.0	55	6	Q46474
34	101	38.8	167	4	P78451
35	101	38.8	200	13	Q8QF48
36	99.5	38.3	145	6	Q9BDR4
37	99	38.1	200	13	Q8QGJ1
38	99	38.1	200	13	Q8QFN2
39	96	36.9	50	13	Q9PRP5
40	95	36.5	200	13	Q8AV82
41	90	34.6	186	13	Q50283
42	87	33.5	168	13	Q8UVE2
43	87	33.5	168	13	Q9CW26
44	87	33.5	168	13	Q90W27
45	87	33.5	168	13	Q98SR7

ALIGNMENTS

RESULT 1

Q6WNEQ PRELIMINARY: PRT: 217 AA.
AC Q6WNEQ
AT Q6WNEQ
CT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
ET 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Growth hormone.
GN GH-N.
GS Ateles geoffroyi (Black-headed spider monkey).
CC Eukaryota; Metazoa; Chordata; Gratiata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCPI_taxid:9509;
RN [1]
RP SEQUENCE FROM N.A.
RA Revol A., Esquivel L., Santiago D., Barrera-Saidana R.:
RT "Independent duplication of the growth hormone gene in three
RT Anthropoid lineages".
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CR EMBL: AF374234; AAL7286.1; -
DR InterPro: IPR001400; Scmatotropin.
DK Pfam: PF06103; hormone.1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA; 24894 MW; 425829FF41EEAAE6 CRC64;

Query Match 95.8%; Score 249; DB 6; Length 217;
Rest local similarity 97.9%; Pred. No. 1.2e-25;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EPTPLSLRFDNMLRAHRLHOLAFTYQEPPEAYIPKEQKYSFLQNP 49

QB 27 EPTPLSLRFDNMLRAHRLHOLAFTYQEPPEAYIPKEQKYSFLQNP 74

RESULT 2

Q14644 PRELIMINARY: PRT: 245 AA.
IC Q14644
AC Q14644;

```

DT 01-JAN-1998 (TRENBLrel. 35, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Placental growth hormone isoform b3H-V4 precursor.
GN HGH-V.
GE
DE Homo sapiens (Human).
CS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCB_TaxID=9606;
OX
[1]
SEQUENCE FROM N.A.
RP
RC TISSUE=Term Placenta;
RX MEDLINE=9433737; PubMed 929596;
RA Boguszewski C.L., Carlsson P.A., Carlsson T., Clark K.,
RA Carlsson L.M.S., Carlsson P.A.
RT "Cloning of two novel, growth hormone transcripts expressed in human
RI placenta."
RL J. Clin. Endocrinol. Metab. 83:2878-2886(1998).
RL EMBL: AF006063; AAB71829.1; .
RL HSSP: P01241; 1A22
DR InterPro: IPR031400; Somatotropin.
DR Pfam: PF01010; Hormone; 1.
DR PROSITE: PS00286; SOMATOTROPIN_1; 1.
KW SIGNAL.
FT SIGNAL
FT SIGNAL
SQ SEQUENCE 245 AA; 27101 MW; 140C7F8C075D8108 CRC64;
Query Match 87.7%; Score 228; 25.4; Length 245;
Best Local Similarity 91.7%; Pred. No. 9, 56-23;
Matches 42; Conservative 2; Mismatches 2; Indels 0; Gaps
QY 42 FT:PLSLRFDNMLRAFLRLQALFDIVYCEFEAYIPKCKYSFLQNP 49
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 27 FT:PLSLRFDNMLRAFLRLQALFDIVYCEFEAYIPKCKYSFLQNP 74
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
QC7368
ID QC7368 PRELIMINARY: PRT; 212 AA.
AC QC7368;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Somatotropin 2 precursor (growth hormone 2) (Fragment).
CS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
ON NCB_TaxID=9544;
OX
[1]
SEQUENCE FROM N.A.
RP
RC TISSUE=Placenta;
RX MEDLINE=9433724; PubMed 9404617;
RA Golos T.G., Durning M., Fisher J.M., Fowler J.D.
RT "Cloning of four growth hormone/chorionic somatomotropin-related
RI complementary deoxyribonucleic acids differentially expressed during
RT pregnancy in the rhesus monkey placenta."
RL Endocrinology 134:1744-1752(1998).
RL EMBL: S16533; AAA18840.1; .
RL HSSP: P01241; 1AX1.
DR InterPro: IPR031400; Somatotropin.
DR Pfam: PF01010; Hormone; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
DR NON_TER
FT SIGNAL
SQ SEQUENCE 212 AA; 24525 MW; 275C91136256FE95 CRC64;
Query Match 81.9%; Score 213; 18.7; Length 212;
Best Local Similarity 78.7%; Pred. No. 8, 76-23;
Matches 37; Conservative 9; Mismatches 1; Indels 0; Gaps
QY 3 FT:PLSLRFDNMLRAHRLRLALFDIVYCEFEAYIPKCKYSFLQNP 43
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 23 FT:PLSLRFDNMLRAHRLRLALFDIVYCEFEAYIPKCKYSFLQNP 69
| | | | | | | | | | | | | | | | | | | | | | | | | |

```



```

Matches 35; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 3 PTPLSLRFDNAMLRAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 49
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 28 PSVPLSRFDNMLCAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 74
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 6
Q14407 PRELIMINARY; PRT: 217 AA.
AC Q14407
DI 01-NOV-1996 (TREMELrel. 01, Created)
DI 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DI 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Chorionic somatomammotropin CS-2 (Chorionic somatomammotropin hormone
  2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9605;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84307277; PubMed=2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Selinus R.B.,
RA Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
  evolution.";
RL Genomics 4:473-497(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91102558; PubMed=1480158;
RA Vnencak-Jones C., Phillips J.A.;
RT "Hot spots for growth hormone gene deletions in homologous regions
  outside of Alu repeats.";
RL Science 250:1745-1748(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: J03971; AA852553.1;
DR EMBL: NC022644; AA822044.1;
DR EMBL: PC035963; AA815963.1;
DR HSSP: P01241; IAZ2.
DR InterPro: IPR01400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA; 24994 MW; 39FAACDDB6B2F951 CRC64;

Query Match 75.8%; Score 137; DB 4; Length 217;
Rest Local Similarity 80.0%; Pred. No. 1,30-18;
Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 TPLSLRFDNAMLRAHRLHOLAFDTYOEFEEAYIPKEOKYSFLON 48
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 29 TVPLSRFDNMLCAHRLHOLAFDTYOEFEEAYIPKEOKYSFLON 73
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 7
Q8WNC9 PRELIMINARY; PRT: 217 AA.
AC Q8WNC9;
DI 01-MAR-2002 (TREMELrel. 20, Created)
DI 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DI 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Growth hormone.

Matches 35; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 3 PTPLSLRFDNAMLRAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 49
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 28 PSVPLSRFDNMLCAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 74
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 8
Q8MT74 PRELIMINARY; PRT: 217 AA.
AC Q8MT74;
DI 01-OCT-2002 (TREMELrel. 22, Created)
DI 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DI 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Growth hormone-like protein 6 precursor.
GN GHLP6.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrix.
CX NCBI_TaxID=5483;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallis G.C., Wallis M.;
RL "Characterisation of the GH gene cluster in a new world monkey, the
  marmoset (Callithrix jacchus)";
DR EMBL: AJ489811; CA234512.1;
DR InterPro: IPR01400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
DR Signal.
DR CHAIN 1 26 POTENTIAL.
DR SIGNAL 27 217 GROWTH HORMONE-LIKE PROTEIN 6.
SQ SEQUENCE 217 AA; 25177 MW; 5ECF148798278F1A CRC64;

Query Match 65.4%; Score 170; DB 6; Length 217;
Rest Local Similarity 68.1%; Pred. No. 6,10-15;
Matches 32; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 3 PTPLSLRFDNAMLRAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 49
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 28 PSVPLSRFDNMLCAHRLHOLAFDTYOEFEEAYIPKEOKYSFLONP 74
    ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 9
Q70615 PRELIMINARY; PRT: 216 AA.
AC Q70615;
DI 01-AUG-1998 (TREMELrel. 07, Created)
DI 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DI 01-MAR-2003 (TREMELrel. 23, Last annotation update)

```

DE Growth hormone precursor.
 OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclerognathia; Muridae; Spalacinae;
 OC Nannospalax.
 CX NCBI_TaxID=30637;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124645; PubMed=9924177;
 RA Lioupias A., Nevo E., Wallis M.;
 RT "Cloning and characterisation of the gene encoding mole rat (Spalax
 ehrenbergi) growth hormone."
 RL J. Mol. Endocrinol. 22:25-36(1999).
 DR EMBL: AJ005819; CAAG6716.1;
 DR HSSP: P01241; IAX;
 DR InterPro: IP001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR08336; SOMATOTROPIN.
 DR PROSITE: PS02266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS02338; SOMATOTROPIN_2; 1.
 KW Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 256 GROWTH HORMONE.
 FT SEQUENCE 216 AA; 24627 MW; EEA88A523BA0ADFE CRC64;
 Query Match 61.7%; Score 160.5; DB 1; Length 216;
 Best Local Similarity 68.1%; Pred. No. 1.2e-13;
 Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
 QY 2 FPTIPLSRFLDNAMLRHQLHQAFTDYQEFEEAYIPKECKYSFLON 48
 DB 27 FPAMPSSLSFANAVLRACHLHQAADTYKEFERAYIPKECKYS-ION 72
 RESULT 10
 QYTV91 PRELIMINARY: PRT; 52 AA.
 AC Q9TV91;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Growth hormone (Fragment).
 GN GH.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 CX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99160448; PubMed=19951423;
 RA Gaetano A.R., Pomp G., Murray S.D., Bowling A.T.;
 RT "Comparative mapping of 16 equine type I genes assigned by somatic
 cell hybrid analysis."
 RL Mamm. Genome 10:271-276(1999).
 DR EMBL: AF097589; AA025992.1;
 DR HSSP: P01241; IGH;
 DR InterPro: IP001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 FT NON_TER 1 52
 FT NON_TER 52 52
 FT SEQUENCE 52 AA; 5835 MW; 20A9E9F9139F996A CRC64;
 Query Match 61.7%; Score 159.5; DB 6; Length 52;
 Best Local Similarity 68.1%; Pred. No. 3.2e-14;
 Matches 32; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 2 FPTIPLSRFLDNAMLRHQLHQAFTDYQEFEEAYIPKECKYSFLON 48
 DB 6 FPAMPSSLSFANAVLRACHLHQAADTYKEFERAYIPKECKYS-ION 51

RESULT 11
 Q8MI73

Q8MI73 PRELIMINARY: PRT; 216 AA.
 AC Q8MI73;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Growth hormone precursor.
 GN GH.
 OS Delphinus delphis (Saddleback dolphin) (Black sea dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 OC Delphinus.
 CX NCBI_TaxID=9722;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 RA Manion D., Wallis O.G., Wallis M.;
 RT "Cloning and characterisation of the GH gene from the common dolphin
 (Delphinus delphis)."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ Databases.
 DR EMBL: AJ492191; CAD37292.1;
 DR InterPro: IP001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR08336; SOMATOTROPIN.
 DR PROSITE: PS02266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 216 GROWTH HORMONE.
 FT SEQUENCE 216 AA; 24509 MW; 1EC46784CCFEC2 CRC64;
 Query Match 61.3%; Score 159.5; DB 6; Length 216;
 Best Local Similarity 68.1%; Pred. No. 1.6e-13;
 Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
 QY 2 FPTIPLSRFLDNAMLRHQLHQAFTDYQEFEEAYIPKECKYSFLON 48
 DB 27 FPAMPSSLSFANAVLRACHLHQAADTYKEFERAYIPKECKYS-ION 72
 RESULT 12
 Q8HYE5 PRELIMINARY: PRT; 216 AA.
 AC Q8HYE5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Growth hormone precursor.
 GN GH.
 OS Alluropoda melanoleuca (Giant panda).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Alluropoda.
 CX NCBI_TaxID=9646;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Pituitary;
 RA Gao M., Zhu M., Zhang A.;
 RT "Cloning and expression of cDNA encoding growth hormone from
 Alluropoda melanoleuca."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ Databases.
 DR EMBL: AF540936; AAN77228.1;
 KW Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT SEQUENCE 216 AA; 24393 MW; 44EC17EC44BCB056 CRC64;
 Query Match 61.3%; Score 159.5; DB 6; Length 216;
 Best Local Similarity 68.1%; Pred. No. 1.6e-13;
 Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
 QY 2 FPTIPLSRFLDNAMLRHQLHQAFTDYQEFEEAYIPKECKYSFLON 48
 DB 27 FPAMPSSLSFANAVLRACHLHQAADTYKEFERAYIPKECKYS-ION 72
 RESULT 13

```

RESULT 15
Q95M75 PRELIMINARY: PRT: 178 AA.
AC Q95M75
DT 01-DEC-2001 (TrEMBLrel: 19, Created)
DT 01-DEC-2001 (TrEMBLrel: 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel: 23, Last annotation update)
DE Growth hormone (Fragment).
OS Tarsius bancanus (western tarsier) (Horsfield's tarsier).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Tarsii; Tarsiidae; Tarsius.
OX NCBI_TaxID=9477;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:21265430; PubMed=11371582;
RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT Species Specificity of Human Growth Hormone Receptor.";
RI Mol. Biol. Evol. 18:945-953(2001).
ER EXBL: AF339081; AAK62338.1; -.
DR InterPro: IPR001400; Somatotropin.
DE Prim; PFG0103; hormone; i.
PR PROSITE: PR00836; SOMATOTROPIN.
ER PROSITE: PS00266; SOMATOTROPIN_1; i.
FT NON-TER 1
FT NON-TER 178
SQ SEQUENCE :78 AA: 20038 MW: 672087512PSB7F0 CRC64;
Query Match: 58.5%; Score 152; DB 6; Length 178;
Best local similarity: 65.1%; Prec. No. 1.3e-12;
Matches 28; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
QY 2 FFTIPSLFLFNAMLRHRLHQLAFDTYQEFEEAYIPKPKYKS 44
DP 11::11 111 1111 1111 111:11 111:11 1
8 FSVSLSLFLFANAVLRACHLHQLAADTYKEFKTHIFEDQNS 5C
Search completed: September 15, 2003, 12:03:27
Job time : 22.1434 secs

```

GenCore version 5.1.1
Copyright (c) 1993 - 2003 Computer Ltd.
OM protein - protein search, using sw model
Run on: September 15, 2003, 11:44:15 : Search time 14.6406 seconds
(without alignments)
225,942 Million cell updates/sec.

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MEPTIPSLRFDNMLRAHR.....NLEFLRSLLLISWELRQ 92
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0
Maximum Match 100%
Listing first 15 summaries

Database : A_Genoseq_10Jun03:*

- 1: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1983.DAT*
- 2: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1984.DAT*
- 3: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1985.DAT*
- 4: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1986.DAT*
- 5: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1987.DAT*
- 6: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1988.DAT*
- 7: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1989.DAT*
- 8: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1990.DAT*
- 9: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1991.DAT*
- 10: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1992.DAT*
- 11: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1993.DAT*
- 12: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1994.DAT*
- 13: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1995.DAT*
- 14: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1996.DAT*
- 15: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1997.DAT*
- 16: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1998.DAT*
- 17: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA1999.DAT*
- 18: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2000.DAT*
- 19: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2001.DAT*
- 20: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2002.DAT*
- 21: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2003.DAT*
- 22: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2004.DAT*
- 23: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2005.DAT*
- 24: /SIDSL/qcqdta/geneseq/geneseq-emb1/AA2006.DAT*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DR	ID	Description
1	470	100.0	92	20	AAV42856	Human growth hormone
2	470	100.0	134	20	AAW92265	Human anti-angiogen
3	470	100.0	150	20	AAV42861	Chimeric protein,
4	465	98.9	140	10	AAV91041	Human growth hormone
5	465	98.9	192	10	AAV90129	Human growth hormone
6	465	98.9	192	20	AAW92264	Human anti-angiogen
7	465	98.9	261	10	AAV92299	Human nerve growth
8	465	98.9	262	12	AAK11740	Human growth hormone
9	465	98.9	310	11	AAW92255	Fusion protein of

10	452	98.3	144	11	AAW93113	Segment of B-cell
11	452	98.3	262	7	AAV61033	Human beta-nerve g
12	450	97.9	138	9	AAV81226	Sequence of protei
13	450	97.9	191	18	AAO20110	Primary amino acid
14	450	97.9	191	20	AAV15809	Protein sequence o
15	450	97.9	191	20	AAV04396	Natural human 22kD
16	450	97.9	191	20	AAV04397	Mutant human 22kDa
17	450	97.9	191	21	AAV78425	Human growth hormo
18	450	97.9	191	22	AAO17485	Human growth hormo
19	450	97.9	191	22	AAO17486	Human growth hormo
20	450	97.9	191	23	ABG94859	Human growth hormo
21	450	97.9	191	23	ABG94860	Human growth hormo
22	450	97.9	191	23	ABG94861	Human growth hormo
23	450	97.9	191	23	ABG94862	Human growth hormo
24	450	97.9	191	23	ABG94863	Human growth hormo
25	450	97.9	191	23	ABG94864	Human growth hormo
26	450	97.9	191	23	ABG94907	Human growth hormo
27	450	97.9	191	23	ABG94908	Human growth hormo
28	450	97.9	191	23	ABG94909	Human growth hormo
29	450	97.9	191	23	ABG94910	Human growth hormo
30	450	97.9	191	23	ABG94911	Human growth hormo
31	450	97.9	191	23	ABG94912	Human growth hormo
32	450	97.9	191	23	ABG94913	Human growth hormo
33	450	97.9	191	23	ABG94914	Human growth hormo
34	450	97.9	191	23	ABG94915	Human growth hormo
35	450	97.9	191	23	ABG94916	Human growth hormo
36	450	97.9	191	23	ABG94917	Human growth hormo
37	450	97.9	191	23	ABG94918	Human growth hormo
38	450	97.9	191	23	ABG94919	Human growth hormo
39	450	97.9	191	23	ABG94920	Human growth hormo
40	450	97.9	191	23	ABG94921	Human growth hormo
41	450	97.9	191	23	ABG94922	Human growth hormo
42	450	97.9	191	23	ABG94923	Human growth hormo
43	450	97.9	191	23	ABG94924	Human growth hormo
44	450	97.9	191	23	ABG94925	Human growth hormo
45	450	97.9	191	23	ABG94926	Human growth hormo

ALIGNMENTS

HEXIDE 1
AAV42856
ID : AAV42856 standard; protein: 92 AA.
AC AAV42856.
XX
XX 15-JAN-2000 (first entry)
XX Human growth hormone (hGH) N-terminal fragment #2.
XX Growth hormone; claptose; intramolecular; insulin; precursor;
XX Folding; conformation; chimeric protein; cleavable; recombinant;
XX Production; yield.
XX
XX Homo sapiens.
XX
XX WC9550302-A1.
XX
XX 07-OCT-1999.
XX
XX 31-MAR-1998; 98WO-CN00052.
XX
XX 31-MAR-1998; 98WO-CN00052.
XX
XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX Gen 2;
XX
XX WPI; 1559-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used
PI particularly for the production of human insulin -

PN W09950302-A1.
 XX 07-OCT-1999.
 XX 31-MAR-1998; 9BWO-CN00052.
 XX 31-MAR-1998; 9BWO-CN00052.
 XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX Gan %;
 XX WPI; 1999-610339/52.
 XX New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
 XX Claim 14; Page 30-31; 46pp; English.
 XX This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin precursor comprising insulin A and B chains (AAY42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (INC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. CC Production of recombinant human insulin via an hGH-proinsulin chimera CC protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The INC sequences not only protect insulin CC sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the CC solubility of the fusion protein and decrease the intermolecular CC interactions among the fusion proteins, thus allowing foldings of the CC fused insulin precursor at commercially useful high concentrations. The CC procedural steps of cyanogen bromide cleavage, oxidative sulphhydrylase and related purification steps can thus be eliminated, along with the use CC of high concentrations of mercaptan or the use of hydrophobic absorbent CC resins.
 XX SQ Sequence 150 AA;
 Query Match 100.0%; Score 470; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 10-39;
 Matches 92; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPTTIPSLRFLFNAMLRHRLHQLAFDTYQEEAEYIPKEQKYSFLQNPQTSLSFSESIP 60
 DB 1 MPTTIPSLRFLFNAMLRHRLHQLAFDTYQEEAEYIPKEQKYSFLQNPQTSLSFSESIP 60
 QY 61 TPSNREETQOKSNLELLRSLILQSWLEPVQ 92
 DB 61 TPSNREETQOKSNLELLRSLILQSWLEPVQ 92
 RESULT 4
 AAP91041
 IC AAP91041 standard; protein, 145 AA.
 XX AAP91041;
 XX 14-DEC-1989 (first entry)
 XX Human growth hormone segment.
 XX Human growth hormone; fusion protein; thrombin;
 XX geriatric dementia; nervous disorders; human nerve factor.
 XX Homo sapiens (human).
 XX EP329175-A.
 XX

PN 23-AUG-1989.
 XX 17-FEB-1989; 89EP-0102795.
 XX 19-FEB-1988; 86JP-0035042.
 XX (ICXJ) TOSOH CORP.
 XX Ohtsuka E;
 XX WPI; 1989-243092/34.
 XX New human nerve growth factor gene encoding fusion protein
 PT - having cleavage site for thrombin, useful for treating geriatric
 PT dementia, etc.
 XX Disclosure; page 21; 38pp; English.
 XX Human growth hormone segment, used at the N-terminal of a fusion
 CC protein, which contains a thrombin recognition site, and human beta nerve
 CC growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to
 CC control geriatric dementia and other nervous disorders, and can be
 CC released from the fusion protein by incubation with thrombin (see
 CC AAN90577-8, AAP91034, AAP91299).
 XX SQ Sequence 140 AA;
 Query Match 98.9%; Score 455; DB 10; Length 140;
 Best Local Similarity 98.9%; Pred. No. 3e-39;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MPTTIPSLRFLFNAMLRHRLHQLAFDTYQEEAEYIPKEQKYSFLQNPQTSLSFSESIP 60
 DB 1 MPTTIPSLRFLFNAMLRHRLHQLAFDTYQEEAEYIPKEQKYSFLQNPQTSLSFSESIP 60
 QY 61 TPSNREETQOKSNLELLRSLILQSWLEPVQ 92
 DB 61 TPSNREETQOKSNLELLRSLILQSWLEPVQ 92
 RESULT 5
 AAP90129
 IC AAP90129 standard; protein, 192 AA.
 XX AAP90129;
 XX 25-MAR-2003 (updated)
 XX 06-FEB-1996 (revised)
 XX 01-NOV-1989 (first entry)
 XX Human growth hormone.
 XX Human growth hormone; fusion protein; recombinant
 XX vector.
 XX Homo sapiens (Human).
 XX JF01144981-A.
 XX 07-JUN-1989.
 XX 02-DEC-1987; 87JP-0304937.
 XX 02-DEC-1987; 87JP-0304937.
 XX (WAKT) WAKUNAGA SEIYAKU KK.
 XX WPI; 1989-209284/29.
 XX N-PSDB; AAN90269.
 XX Recombinant vector config. fusion protein - consisting of human
 PT growth hormone or deriv. ligated to foreign protein, for stability
 PT and high yield.
 PT

XX Disclosure: Fig 1: 19pp; Japanese.
PS
CC The invention consists of a vector contg. a fusion protein which is
CC formed by ligating downstream of a promoter, hGH or a deriv. (prot.
CC formed by substn. of Met-14 with leu) and a foreign protein.
CC Stability of the vector in the host is greatly increased so the
CC protein yield is higher.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 192 AA:
XX
XX Query Match 98.9%; Score 465; DB 16; Length 192;
XX Best Local Similarity 98.9%; Pred. No. 4,2e-39;
XX Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MFTPTLSRLFDNMLRAHRLHQLAFDYQFEFAYIPKEQKYSFLONPQTSLSFSESIP 60
XX DQ 1 MFTPTLSRLFDNMLRAHRLHQLAFDYQFEFAYIPKEQKYSFLONPQTSLSFSESIP 60
XX QY 61 TPSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
XX DQ 61 TPSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
XX
XX RESULT 6
XX AAW92264
XX ID AAW92264 standard; Protein: 192 AA.
XX AC AAW92264;
XX DT 08-JUN-1999 (first entry)
XX DE Human anti-angiogenic peptide hGH Met-1phe191.
XX
XX Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
XX growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
XX placental vasculature; pregnancy; treatment; angiogenic disease;
XX tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;
XX arthritis; atherosclerotic plaques; corneal graft neovascularisation;
XX wound healing; proliferative retinopathy; macular degeneration; trachoma;
XX granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;
XX psoriasis; fibropia; scleroderma; Kaposi's sarcoma; vascular adhesion;
XX ulcer; leukaemia; reproductive disorder; contraceptive agent;
XX gene therapy; pre-eclampsia; intrauterine growth retardation;
XX placental dysfunction.
XX
XX Homo sapiens.
XX
XX WO9851323-A1.
XX
XX 19-NOV-1998.
XX
XX 12-MAY-1998; 98WO-US09491.
XX
XX 13-MAY-1997; 97US-0046394.
XX
XX (RECC) UNIV CALIFORNIA.
XX
XX Martial JA, Struman J, Taylor E, Welner RJ;
XX
XX WPI: 1999-045192/04.
XX N-PSDB: AAX01706.
XX
XX New anti-angiogenic peptides - comprise N-terminal fragments of
XX human placental lactogen, human growth hormone, growth hormone
XX variant or human prolactin.
XX
XX Example 3; Page 49; 87pp; English.
XX
XX This invention describes novel human anti-angiogenic peptides derived
XX from 10 to 150 consecutive amino acids selected from the N-terminal end
XX of human placental lactogen (hPL), human growth hormone (hGH), growth

XX hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
XX capillary endothelial cell proliferation and organisation (ii) inhibit
XX angiogenesis in chick chorioallantoic membrane and (iii) binds to at
XX least one specific receptor which does not bind an intact full length
XX hGH, hPL, prolactin or hGH-V. The invention also describes a method for
XX diagnosing a probable abnormality of the placental vasculature during
XX pregnancy. The peptides can be used for treating an angiogenic disease in
XX a subject, for inhibiting tumour formation or growth in a patient or for
XX modulating vascularisation of a patient's placenta. In particular, the
XX peptides can be used for preventing or treating e.g. malignant tumours,
XX angiofibroma, arteriovenous malformation, arthritic such as rheumatoid
XX arthritis, atherosclerotic plaques, corneal graft neovascularisation,
XX retinopathy, macular degeneration, granulations such as diabetic
XX such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular
XX tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis,
XX pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,
XX Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,
XX leukaemia, and reproductive disorders such as follicular and luteal cysts
XX and choriocarcinoma. They can also be used as contraceptive agents, DNA
XX encoding the peptides can be used in gene therapy. The measurement of
XX abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL
XX can be used in assays for impairment of vascular development associated
XX with pre-eclampsia, intrauterine growth retardation, and p-acental
XX dysfunction.
XX
XX Sequence 192 AA:
XX
XX Query Match 98.9%; Score 465; DB 20; Length 192;
XX Best Local Similarity 98.9%; Pred. No. 4,2e-39;
XX Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MFTPTLSRLFDNMLRAHRLHQLAFDYQFEFAYIPKEQKYSFLONPQTSLSFSESIP 60
XX DQ 1 MFTPTLSRLFDNMLRAHRLHQLAFDYQFEFAYIPKEQKYSFLONPQTSLSFSESIP 60
XX QY 61 TPSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
XX DQ 61 TPSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
XX
XX RESULT 7
XX AAW91299
XX ID AAW91299 standard; Protein: 261 AA.
XX AC AAW91299;
XX DT 14-DEC-1989 (first entry).
XX
XX Human nerve growth factor and human growth hormone fusion protein.
XX
XX Human nerve growth factor; fusion protein; thrombin;
XX geriatric dementia; nervous disorders; human growth hormone.
XX
XX Homo sapiens (human).
XX
XX Key Location/Qualifiers
XX Region 1..140
XX Region 141..143
XX Region 144..261
XX
XX EF329175-A.
XX
XX 23-AUG-1989.
XX
XX 17-FEB-1989; 89EP-0102795.
XX
XX 19-FEB-1986; 89JP-0035042.
XX
XX (TQVJ) TOSOH CORP.
XX
XX Ohtsuka E;
XX
XX


```

SQ Sequence 135 AA.
Query Match 97.9%; Score 460; DB 9; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLRISILLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLRISILLIQSWLEPVQ 91

RESULT 13
ID AAO20110 standard; protein: 191 AA.
AC AAO20110;
XX
XX
DT 06-AUG-2002 (first entry)
DE
DE Protein sequence of the hGH growth hormone cDNA.
DE Serum albumin-growth hormone fusion protein; growth hormones;
KW Down's syndrome.
XX
XX Unidentified.
XX
XX KR99076789-A.
XX
XX 15-OCT-1999.
XX
XX 25-JUN-1998; 96KR-070491A.
XX
XX 30-DEC-1995; 95GB-002672A.
XX
XX 19-DEC-1996; 96WO-GH0316A.
XX
XX (DELTA ) DELTA BIOTECHNOLOGY LTD.
XX
XX WPI: 1997-36369C/55.
XX
XX N-PSDB: AAX99565.
XX
XX Serum albumin-growth hormone fusion protein - useful to treat growth
PT hormone related diseases, e.g. Down's syndrome
XX
XX Disclosure: Fig 1; 2:pp; Karman.
XX
XX The invention relates to a serum albumin-growth hormone fusion protein
CC useful to treat growth hormone related diseases such as Down's syndrome.
CC This sequence represents a protein of the serum albumin-growth hormone
CC cDNA of the invention.
XX
XX Sequence 191 AA:

Query Match 97.9%; Score 460; DB 10; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLRISILLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLRISILLIQSWLEPVQ 91

RESULT 14
AAV15809
ID AAV15809 standard; protein: 191 AA.
XX

```

```

AC AAV15809;
XX
XX 28-JUL-1999 (first entry)
DE
DE Primary amine acid sequence of native human growth hormone.
XX
XX detection; fluoresce; illeal misuse; growth substance; athlete;
KW domesticated farm animal; cattle; human growth hormone.
XX
XX Homo sapiens.
XX
XX WC9926069-A1.
XX
XX 27-MAY-1999.
XX
XX 15-NOV-1998; 98WO-GH03449.
XX
XX 14-NOV-1997; 97GB-G023955.
XX
XX (GENE-) GENERIC BIOLOGICALS LTD.
XX
XX Atkinson A, Murphy JP;
XX
XX WPI: 1999-336072/28.
XX
XX Use of tagged exogenous polypeptide
XX
XX Disclosure: Fig 1; 36pp; English.
XX
XX The specification describes a method of detecting an exogenously
CC administered substance from a naturally-occurring endogenous substance,
CC the exogenous substance being tagged so that it fluoresces differently
CC from the endogenous one at a suitable wavelength. The tagging may
CC consist of one or more substitutions in tagged growth hormone
CC selected from G40Y, F52Y, W86F, Y, L, I or V I103Y or I137Y;
CC The method is used to distinguish between exogenously administered
CC substances as compared to naturally-occurring endogenous substances.
CC Especially mentioned is the illegal misuse of growth substances by
CC athletes or in domesticated farm animals e.g. cattle. The present
CC sequence represents native human growth hormone which may be used
CC in the method of the invention.
XX
XX Sequence 191 AA:

Query Match 97.9%; Score 460; DB 20; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTPLSRFLDNAMLRARHLQHLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLRISILLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLRISILLIQSWLEPVQ 91

RESULT 15
AAV04396
ID AAV04396 standard; protein: 191 AA.
XX
XX AAV04396;
XX
XX 29-JUN-1999 (first entry)
XX
XX Natural human 22kda growth hormone.
XX
XX Human; 22kda growth hormone; hGH; mutant; thrombin; resistance;
KW plasmin; decomposition.
XX
XX Homo sapiens.
XX
XX JPI1092499-A.
XX

```

```

XX 06-APR-1999.
PD
XX
XX 22-SEP-1997; 97JP-0275277.
PF
XX
XX 22-SEP-1997; 97JP-0275277.
PR
XX
XX (SUMO ) SUMITOMO SEIYAKU KK.
PA
XX
XX WPI; 1999-283567/24.
DR
XX
XX A human growth hormone mutant - with equivalent activity to natural
PI human growth hormone
XX
XX Example 1; Page 5-6; 10pp; Japanese.
PS
XX
XX The present invention describes a human growth hormone mutant in which
CC the 134th Arg and the 155th Thr are replaced respectively by Asp and Phe
CC in the 1st to the 191st amino acid sequence of natural type human 22 kDa
CC growth hormone (hGH) and which has a resistance against decomposition by
CC thrombin. The present sequence represents the natural hGH. Also
CC described are: (1) a hGH mutant in which the 134th Arg, the 155th Thr
CC and the 140th Lys are replaced respectively by Asp, Pro and Ala in the
CC amino acid sequence of natural type hGH and which has a resistance
CC against decomposition by thrombin and plasmin; and (2) a drug
CC preparation containing the above hGH mutant as the active component.
CC The mutant hGH shows an activity approximately equivalent to that of
CC natural type hGH and shows a high stability in blood and body fluid.
XX
XX Sequence 19) AA:
SQ
Query Match 97.9%; Score 460; DB 20; Length 191;
Best Local Similarity 99.9%; Pred. NC, 1,30-167;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 FTPIPLSLFDNANLPAHSLHQLADTYGEFEAVIPKCKYSPFLGNQVLSLSPSSSPT 60
DB 1 FTPIPLSLFDNANLPAHSLHQLADTYGEFEAVIPKCKYSPFLGNQVLSLSPSSSPT 60
QY 62 PSNRRTCKSKNLELRISLQLTQSWLEPVQ 92
DB 61 PSNRRTCKSKNLELRISLQLTQSWLEPVQ 91

```

Search completed: September 15, 2003, 12:00:57
 Job time : 65.6308 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein : protein search, using sw model

Run on: September 15, 2003, 11:52:40 : Search time 21,109 seconds
(without alignments)
184,449 Million cell updates/sec

Title: US-09-423-100-2

Perfect score: 472

Sequence: 1 MFPIPLSRFDNAKRAHR.....NELLRISLLIQSRIHWQ 92

Scoring table: BLDSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4230658 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2500000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents_AA*

1: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*
2: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*
3: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*
4: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*
5: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*
6: /cgn2_6/plodata/1/iaa/5A_COMB.rep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	98.9	192	US-08-293-383-1	Sequence 1, Appli
2	460	97.9	191	US-09-284-878-4	Sequence 5, Appli
3	460	97.9	194	US-08-383-621-4	Sequence 5, Appli
4	460	97.9	194	US-08-459-508-4	Sequence 5, Appli
5	460	97.9	217	US-08-589-038-10	Sequence 1, Appli
6	460	97.9	217	US-08-784-582-10	Sequence 1, Appli
7	460	97.9	217	US-08-788-271-10	Sequence 1, Appli
8	460	97.9	217	US-08-759-628-11	Sequence 1, Appli
9	460	97.9	217	US-09-284-878-1	Sequence 1, Appli
10	460	97.9	241	US-09-424-6205-25	Sequence 28, Appli
11	460	97.9	245	US-09-285-030-56	Sequence 66, Appli
12	460	97.9	274	US-08-784-582-71	Sequence 71, Appli
13	460	97.9	360	US-08-784-582-71	Sequence 71, Appli
14	454	96.6	191	US-09-465-461-1	Sequence 1, Appli
15	454	96.6	217	US-08-187-756C-4	Sequence 1, Appli
16	454	96.6	217	US-08-459-486-51	Sequence 51, Appli
17	454	96.6	217	US-08-459-658-51	Sequence 51, Appli
18	454	96.6	217	US-08-710-324A-4	Sequence 4, Appli
19	454	96.6	217	US-09-411-657-4	Sequence 4, Appli
20	453	96.4	400	US-09-420-819-37	Sequence 37, Appli
21	453	96.4	401	US-09-420-819-36	Sequence 36, Appli
22	447	95.1	191	US-08-800-25C-18	Sequence 18, Appli
23	445	94.7	191	US-08-800-25C-26	Sequence 16, Appli
24	445	94.7	191	US-08-800-25C-20	Sequence 20, Appli
25	364.5	77.6	176	US-08-791-728-1	Sequence 1, Appli
26	364.5	77.6	176	US-08-590-774-1	Sequence 1, Appli
27	358.5	76.3	176	US-08-791-728-2	Sequence 2, Appli

ALIGNMENTS

RESULT: 1

US-08-093-383-1

: Sequence 1, Application US/08093383

: Patent No. 5489529

: GENERAL INFORMATION:

: APPLICANT: DeBoer, Herman A.

: APPLICANT: Heyneker, Herbert L.

: APPLICANT: Seeburg, Peter H.

: TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone

: NUMBER OF SEQUENCES: 30

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.

: STREET: 460 Point San Bruno Blvd

: CITY: South San Francisco

: STATE: California

: COUNTRY: USA

: ZIP: 94080

: COMPUTER READABLE FORM:

: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: palin (Genentech)

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/093.383

: FILING DATE: 14-JUL-1993

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 07/619827

: FILING DATE: 28-NOV-1990

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 07/198824

: FILING DATE: 05-APR-1986

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 06/632361

: FILING DATE: 19-JUL-1984

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 06/303687

: FILING DATE: 16-SEP-1981

: ATTORNEY/AGENT INFORMATION:

: NAME: Johnston, Sean A.

: REGISTRATION NUMBER: PJ5,910

: REFERENCE/DOCKET NUMBER: 46C4

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 415/225-3562

: TELEFAX: 415/952-9881

: TELEX: 910/371-7168

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 192 amino acids

: TYPE: amio acid

Sequence 2, Appli
Patent No. 5424199
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 8, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Patent No. 5210180
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli

28 358.5 76.4 174 4 US-08-990-774-2
29 340 72.3 168 6 5424199-3
30 323.5 71.2 158 1 US-08-187-756C-5
31 358.5 71.2 158 2 US-08-710-324A-5
32 343.5 71.2 158 4 US-09-411-657-5
33 306.5 65.2 191 1 US-08-468-824-8
34 304.5 64.9 191 1 US-07-963-331D-4
35 302.5 64.4 216 2 US-09-105-651-1
36 301.5 64.1 190 1 US-08-388-267C-2
37 301.5 64.1 190 4 US-09-277-720-2
38 301.5 64.1 191 6 5210180-1
39 301.5 64.1 193 1 US-07-621-197C-2
40 301.5 64.1 193 1 US-08-363-982-2
41 301.5 64.1 193 2 US-08-383-621-1
42 301.5 64.1 193 3 US-08-459-906-1
43 301.5 64.1 216 2 US-09-105-651-3
44 299.5 63.7 190 1 US-07-963-331D-3
45 299.5 63.7 191 1 US-07-922-523-1


```

? REGISTRATION NUMBER: 34,276
? REFERENCE/DOCKET NUMBER: 31,278-53
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-831-3247
? TELEFAX: 201-831-3305
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 194 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-459-906-4

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 7,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 63

QY 62 PSNREETOOKSNLELLRISLLIQT.SWLEPVQ 92
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 64 PSNREETOOKSNLELLRISLLIQT.SWLEPVQ 94

RESULT 5
US-08-589-028-10
? Sequence 10, Application US/C8589028
? Patent No. 6087129
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher B.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 6087129minston, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Aimee E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P. O. Box 4433
? CITY: Houston
? STATE: TX
? COUNTRY: USA
? ZIP: 77210-4433
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/589,028
? FILING DATE: Concurrently Herewith
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/028,427
? FILING DATE: 15-OCT-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/589,028
? FILING DATE: 19-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Highlander, Steven J.
? REGISTRATION NUMBER: 37,642
? REFERENCE/DOCKET NUMBER: UTSD:514
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 512/418-3000
? TELEFAX: 512/474-7577
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 217 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? US-08-589-028-10

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 7,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 86

US-08-589-028-10
? Sequence 10, Application US/C8784582
? Patent No. 6110707
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher B.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 6110707minston, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Aimee E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? NUMBER OF SEQUENCES: 79
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P. O. Box 4433
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77210
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/784,582
? FILING DATE: Concurrently Herewith
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/028,427
? FILING DATE: 15-OCT-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/589,028
? FILING DATE: 19-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Highlander, Steven J.
? REGISTRATION NUMBER: 37,642
? REFERENCE/DOCKET NUMBER: UTSD:514
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 512/418-3000
? TELEFAX: 512/474-7577
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 217 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? US-08-784-582-10

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 7,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27 PPTTPI.SKLF.NAML.RAHL.HQLA.FDTYQE.FEAYIPKEQKYS.FLQNPQT.SLSE.SIPT 86

```

Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117

RESULT 7

US-08-785-271-10
 : Sequence 10, Application US/23785271
 : Patent No. 6194176
 : GENERAL INFORMATION:
 : APPLICANT: Newgard, Christopher H.
 : APPLICANT: Haibian, Philippe A.
 : APPLICANT: No. 5194176 Inventor: Karl D.
 : APPLICANT: Clark, Samuel A.
 : APPLICANT: Thigpen, Anice E.
 : APPLICANT: Quade, Christan
 : APPLICANT: Kruse, Fred
 : TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 : TITLE OF INVENTION: SECRETORY CELL LINES
 : NUMBER OF SEQUENCES: 56
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Arnold, White & Burke
 : STREET: P.O. Box 4433
 : CITY: Houston
 : STATE: Texas
 : COUNTRY: USA
 : ZIP: 77210
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/785,271
 : FILING DATE: Concurrently Herewith
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/589,028
 : FILING DATE: 19-09-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Highlander, Steven L.
 : REGISTRATION NUMBER: 37,542
 : REFERENCE/DOCKET NUMBER: 0758513
 : TELEPHONE: 512/418-3000
 : TELEFAX: 512/474-7577
 : INFORMATION FOR SEQ ID NO: 10:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 217 amino acids
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: linear
 : FEATURE:
 : US-08-785-271-10

Query Match 97.9%; Score 460; DB 2; Length 217;
 Best Local Similarity 98.9%; Pred. No. 7.1e-51;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 FPTIPLSRLEFDNMLRAHLRLHQLAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 61
 Db 27 FPTIPLSRLEFDNMLRAHLRLHQLAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 86
 Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117
 RESULT 8
 US-08-759-628-11
 : Sequence 11, Application US/28759628
 : Patent No. 6225446
 : GENERAL INFORMATION:
 : APPLICANT: Altman, Scott W.

APPLICANT: Rock, Fernando L.
 APPLICANT: Bazan, J. Fernando
 APPLICANT: Kasteleiner, Robert A.
 TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DNAX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/759,628
 FILING DATE: 05-DEC-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/008,574
 FILING DATE: 06-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090
 REFERENCE/DOCKET NUMBER: 0X05520
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-852-9298
 TELEFAX: 415-496-1200
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 32..53
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 94..115
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 133..153
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 192..210
 OTHER INFORMATION: /note= *The peptides above are
 OTHER INFORMATION: depicted in Figure 1*
 US-08-759-628-11

Query Match 97.9%; Score 460; DB 3; Length 217;
 Best Local Similarity 98.9%; Pred. No. 7.1e-51;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 FPTIPLSRLEFDNMLRAHLRLHQLAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 61
 Db 27 FPTIPLSRLEFDNMLRAHLRLHQLAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 86
 Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117

RESULT 9
 US-09-284-878-1
 : Sequence 1, Application US/09284878
 : Patent No. 6342375
 : GENERAL INFORMATION:
 : APPLICANT: Olazaran, Martha Guerrero

```

? APPLICANT: Seidara, Hugo Barrera
? APPLICANT: Salvado, Jose Maria Viador
? TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
? TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
? FILE REFERENCE: 1829.5310000
? CURRENT APPLICATION NUMBER: US/09/244,878
? PRIOR APPLICATION NUMBER: 1999-07-21
? PRIOR FILING DATE: 1997-10-24
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 1
? LENGTH: 217
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-284-878-1

Query Match 97.9%; Score 460; DB 4; Length 217;
Best Local Similarity 98.9%; Pred. No. 7.1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 27 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 61
QY 62 PSNREETOQKSNLELRISLLLIQSWLEPVQ 92
Db 87 PSNREETOQKSNLELRISLLLIQSWLEPVQ 117

RESULT 10
US-09-424-6203-25
? Sequence 25, Application US/994246203
? Patent No. 5391595
? GENERAL INFORMATION:
? APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
? JANG, Ki-Ryong
? MOON, Jae-Woong
? BAE, Cheon-Soon
? YANG, Doo-Suk
? LEE, Jee-Won
? SEONG, Baek-Lin
? TITLE OF INVENTION: Process for preparing recombinant proteins using highly
? efficient expression vector from Saccharomyces cerevisiae
? NUMBER OF SEQUENCES: 25
? CORRESPONDENCE ADDRESS:
? ADDRESSER: BACEMAN S LAPINTE, P.C.
? STREET: Suite 1261, 950 Chapel Street
? CITY: New Haven
? STATE: Connecticut
? COUNTRY: U.S.A.
? ZIP: 06510-2802
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
? COMPUTER: IBM
? OPERATING SYSTEM: WINDOWS 95/98
? SOFTWARE: MS WORD
? CURRENT APPLICATION DATA: US/09/424,6203
? FILING DATE: 24-MO. 5991595-1999
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 241 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: PROTEIN
? SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-424-620B-25

Query Match 97.9%; Score 460; DB 4; Length 241;
Best Local Similarity 98.9%; Pred. No. 8.2e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 51 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 110
QY 62 PSNREETOQKSNLELRISLLLIQSWLEPVQ 92
Db 111 PSNREETOQKSNLELRISLLLIQSWLEPVQ 141

RESULT 11
US-09-280-030-66
? Sequence 66, Application US/09280030A
? Patent No. 6506595
? GENERAL INFORMATION:
? APPLICANT: Sato, Seiji
? APPLICANT: Higashikuni, Naohiko
? APPLICANT: Kudo, Toshiyuki
? APPLICANT: Kondo, Masaki
? TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
? TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
? TITLE OF INVENTION: DNAS
? FILE REFERENCE: 382.1026
? CURRENT APPLICATION NUMBER: US/09/280.030A
? CURRENT FILING DATE: 1999-03-26
? EARLIER APPLICATION NUMBER: JP10-87339/1998
? EARLIER FILING DATE: 1998-03-31
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 66
? LENGTH: 245
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Designated is
? OTHER INFORMATION: an amino acid sequence of MWpMp20-TEV-G-GH
US-09-280-030-66

Query Match 97.9%; Score 460; DB 4; Length 245;
Best Local Similarity 98.9%; Pred. No. 8.4e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 55 FPTPLSLFDNMLRAHRLHQLAFDTYOEFEAYIPKEQKYSFLONPQTSLSFSESPT 114
QY 62 PSNREETOQKSNLELRISLLLIQSWLEPVQ 92
Db 115 PSNREETOQKSNLELRISLLLIQSWLEPVQ 145

RESULT 12
US-08-784-582-71
? Sequence 71, Application US/08784582
? Patent No. 6110707
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher H.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 5110707minoton, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Anice E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? APPLICANT: McGarry, Dennis
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? TITLE OF INVENTION: SECRETORY CELL LINES
? NUMBER OF SEQUENCES: 79
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P.O. Box 4433
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77210

```



```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  FILING DATE: 15-OCT-1995
  APPLICATION NUMBER: US/08/589,028
  FILING DATE: 19-JAN-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Highlander, Steven L.
    REGISTRATION NUMBER: 37,642
    REFERENCE/DOCKET NUMBER: UTSD:514
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 512/418-3000
      TELEFAX: 512/474-7577
  INFORMATION FOR SEQ ID NO: 71:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 214 amino acids
      TYPE: amino acid
      STRANDEDNESS:
      TOPOLOGY: linear
US-08-784-582-71

Query Match          97.9%; Score 460; DP 71; Length 274;
Best Local Similarity 98.9%; Pred. No. 9.9e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 61
DB 27 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 64

QY 62 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 92
DB 87 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 117

RESULT 13
US-08-784-582-73
; Sequence 73, Application US/03784582
; Patent No. 6110707
; GENERAL INFORMATION:
; APPLICANT: Neward, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6110707/Minton, Karl D.
; APPLICANT: Clatk, Samuel A.
; APPLICANT: Thigpen, Anico E.
; APPLICANT: Quade, Christian
; APPLICANT: Kruse, Fred
; APPLICANT: McGarry, Dennis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  FILING DATE: 15-OCT-1995
  APPLICATION NUMBER: US/08/589,028
  FILING DATE: 19-JAN-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Highlander, Steven L.
    REGISTRATION NUMBER: 37,642
    REFERENCE/DOCKET NUMBER: UTSD:514
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 512/418-3000
      TELEFAX: 512/474-7577
  INFORMATION FOR SEQ ID NO: 71:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 214 amino acids
      TYPE: amino acid
      STRANDEDNESS:
      TOPOLOGY: linear
US-08-784-582-71
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  FILING DATE: 15-OCT-1995
  APPLICATION NUMBER: US/09/589,028
  FILING DATE: 19-JAN-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Highlander, Steven L.
    REGISTRATION NUMBER: 37,642
    REFERENCE/DOCKET NUMBER: UTSD:514
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 512/418-3000
      TELEFAX: 512/474-7577
  INFORMATION FOR SEQ ID NO: 73:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 350 amino acids
      TYPE: amino acid
      STRANDEDNESS:
      TOPOLOGY: linear
US-09-784-582-73

Query Match          97.9%; Score 460; DB 3; Length 360;
Best Local Similarity 98.9%; Pred. No. 1.4e-50;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 61
DB 27 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 86

QY 62 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 92
DB 87 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 117

RESULT 14
US-09-465-461-1
; Sequence 1, Application US/09465461
; Patent No. 6348444
; GENERAL INFORMATION:
; APPLICANT: CHAPPEL, Scott
; TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and immune
; FILE REFERENCE: after hematopoietic stem cell transplantation in humans
; CURRENT APPLICATION NUMBER: US/09/465,461
; PRIOR FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 60/112,668
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 191
; TYPE: PRI
; ORGANISM: homo sapiens
US-09-465-461-1

Query Match          96.6%; Score 454; DB 4; Length 191;
Best Local Similarity 97.8%; Pred. No. 3.5e-50;
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 61
DB 1 FPTIPLSRFDNAMLRAHLHQLAFDTYOEFEEAYIPKECKYSFLONPOTSLSFSESIPT 60

QY 62 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 92
DB 61 PSNREETOOKSNLELLRISLLIQTOSWLEPVQ 91

RESULT 15
US-08-187-756C-4
; Sequence 4, Application US/08187756C
```

```

1 Patent No. 5597769
2 GENERAL INFORMATION:
3 APPLICANT: ROSEN, ET AL.
4 TITLE OF INVENTION: Human Growth Hormone
5 NUMBER OF SEQUENCES: 7
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: CARELIA, BYRNE, GAIN, SIFILLIAN,
8 ADDRESSEE: GEOCHI, STEWART & CLSTEIN
9 STREET: 6 BECKER FARM ROAD
10 CITY: ROSELAND
11 STATE: NEW JERSEY
12 COUNTRY: USA
13 ZIP: 07068
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: 3.5 INCH DISKETTE
16 COMPUTER: IBM PS/2
17 OPERATING SYSTEM: MS-DOS
18 SOFTWARE: WORD PERFECT 5.1
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/187,756C
21 FILING DATE: January 27, 1994
22 CLASSIFICATION: 435
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER:
25 FILING DATE:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: FERRARO, GREGORY D.
28 REGISTRATION NUMBER: 36,134
29 REFERENCE/DOCKET NUMBER: 325800-55
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 201-994-1700
32 TELEFAX: 201-594-1744
33 INFORMATION FOR SEQ ID NO: 4:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 217 AMINO ACIDS
36 TYPE: AMINO ACID
37 STRANDEDNESS:
38 TOPOLOGY: LINEAR
39 MOLECULE TYPE: PROTEIN
40 US-08-187-756C.4

```

```

Query Match      96.6%  Score 454  DB 11  Length 217:
Best Local Similarity 97.8%  Pred. No. 4,20-50:
Matches 89: Conservative 6; Mismatches 2; Gaps 0;
QY      2  PPTPLSLRFLNANLPAHRLQLAPTYCEFEERATTPKQKYSFLQNFQTSLSSESLET 61
DB      22  PPTPLSLRFLNANLPAHRLQLAPTYCEFEERATTPKQKYSFLQNFQTSLSSESLET 60
QY      62  PSNREETQCKSNLELRSLILGQSMLEPVC 92
DB      87  PSNREETQCKSNLELRSLILGQSMLEPVC 117

```

Search completed: September 15, 2003, 12:05:31
 Job time : 22.1039 secs

GenCore version 3.1.6
Copyright (c) 1993-2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 15, 2003, 12:03:35 Search time 36,540.6 seconds/
without alignment
347,945 Million cell of hits/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MFPTPLSRFDNMLKAKR.....NLRKSTSLKIDSWLIVD 42

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 541936 seqs, 14,912,426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	470	100.0	92	16	US-10-054-873-2
2	470	100.0	130	12	US-10-054-873-2
3	450	97.9	191	13	US-09-984-010-23
4	463	97.9	191	12	US-10-153-207-1
5	460	97.9	191	13	US-10-400-377-1
6	460	97.9	191	12	US-10-400-708-1
7	460	97.9	214	12	US-10-153-207-6
8	460	97.9	217	9	US-09-929-318-9
9	450	97.9	245	9	US-09-280-030-56
10	454	96.6	217	9	US-09-853-688-2
11	454	95.6	217	14	US-09-943-748C-4
12	453	95.4	217	10	US-09-804-409A-16
13	447	95.1	217	9	US-09-853-688-4
14	444	94.5	191	12	US-09-824-200-12
15	396	84.3	217	9	US-09-850-887-3

381	81.1	163	15	US-10-043-487-350	Sequence 350, Appl
381	81.1	193	12	US-10-153-207-2	Sequence 2, Appl.1
373	73.4	229	15	US-10-103-313-41	Sequence 411, App
345	73.4	246	15	US-10-188-246-16	Sequence 18, Appl
301.5	64.1	190	12	US-10-153-207-3	Sequence 3, Appl1
260	55.3	49	14	US-10-054-873-1	Sequence 1, Appl1
260	55.3	107	14	US-10-054-873-6	Sequence 6, Appl1
174.5	37.1	188	15	US-10-191-879-19	Sequence 19, Appl
174.5	37.1	210	15	US-10-191-879-10	Sequence 10, Appl
162	34.5	146	15	US-10-043-487-337	Sequence 337, App
161.5	34.4	366	10	US-09-887-569A-2	Sequence 2, Appl1
161.5	34.4	388	12	US-10-322-746-4	Sequence 4, Appl1
142	30.2	54	9	US-09-876-478-14	Sequence 14, Appl
132	27.9	106	15	US-10-140-293-32	Sequence 32, Appl
126	26.6	199	15	US-10-140-293-33	Sequence 33, App
125	26.6	199	15	US-10-140-293-25	Sequence 26, App
125	26.6	199	15	US-10-140-293-27	Sequence 27, App
124	26.4	199	15	US-10-140-293-31	Sequence 31, Appl
123	26.2	199	15	US-10-140-293-22	Sequence 22, Appl
122	26.0	199	15	US-10-140-293-21	Sequence 21, Appl
121.5	25.9	125	14	US-10-036-869-25	Sequence 25, Appl
121.5	25.9	206	15	US-10-191-879-2	Sequence 2, Appl1
121.5	25.9	253	14	US-10-036-869-27	Sequence 27, Appl
119	25.3	199	15	US-10-140-293-24	Sequence 24, Appl
118	25.1	199	15	US-10-140-293-18	Sequence 18, Appl
118	25.1	199	15	US-10-140-293-19	Sequence 19, Appl
117	24.9	197	15	US-10-140-293-15	Sequence 15, Appl
116	24.7	197	15	US-10-140-293-13	Sequence 13, Appl
116	24.7	199	15	US-10-153-207-4	Sequence 4, Appl1
116	24.7	199	15	US-10-140-293-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-2
Sequence 2, Application US/10054873
Publication NO. US2002016472A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Biscadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/054,873
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J.
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid

```

: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-054-873-2

Query Match      100.0%; Score 470; DB 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.2e-46;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTPLSLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
DB 1 MPTPLSLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92

RESULT 2
US-10-054-873-7
: Sequence 7, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Ru
: TITLE OF INVENTION: Chimero Protein Containing an
: Intramolecular Clapophone like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/000352
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 21-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 46,946
: REFERENCE/DOCKET NUMBER: 020167-0001300S
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match      100.0%; Score 470; DB 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.1e-46;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTPLSLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
DB 1 MPTPLSLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92

```

```

DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92

RESULT 3
US-09-984-010-23
: Sequence 23, Application US/09984010
: Publication No. US20030104578A1
: GENERAL INFORMATION:
: APPLICANT: Ballanco, David James
: TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
: STREET: 1300 I Street, NW
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/984,010
: FILING DATE: 21-May-2002
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/091,873
: FILING DATE: 25-JUN-1998
: APPLICATION NUMBER: PCT/GR96/03164
: FILING DATE: 19-DEC-1996
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 191 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-984-010-23

Query Match      97.9%; Score 460; DB 11; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.8e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTPLSLKLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 61
DB 1 PPTPLSLKLFNFAMLAHRLHQLAFDTYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 62 PSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 PSNRETOOKSNLELRISLLIOSWLEPVQ 91

RESULT 4
US-10-153-207-1
: Sequence 1, Application US/10153207
: Publication No. US20030153003A1
: GENERAL INFORMATION:
: APPLICANT: James A. Wells
: TITLE OF INVENTION: GROWTH HORMONE VARIANTS
: FILE REFERENCE: 669.12-US-C7
: CURRENT APPLICATION NUMBER: US/10/153,207
: CURRENT FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: 08/479,884
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 08/190,723
: PRIOR FILING DATE: 1994-02-02
: PRIOR APPLICATION NUMBER: 07/960,227

```

```

; PRIOR FILING DATE: 1992-10-13
; PRIOR APPLICATION NUMBER: 07/635,204
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/424,066
; PRIOR FILING DATE: 1989-10-25
; PRIOR APPLICATION NUMBER: 07/264,611
; PRIOR FILING DATE: 1988-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-400-708-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 91

RESULT 5
US-10-400-377-1
; Sequence 1, Application US/10400377
; Publication No. US20030162949A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; APPLICANT: Bolder Biotechnology, Inc.
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4:52-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,377
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-400-377-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 91

RESULT 6
US-10-400-708-1
; Sequence 1, Application US/10400708
; Publication No. US20030166865A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; APPLICANT: Bolder Biotechnology, Inc.
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4:52-1-PUS
```

```

; CURRENT APPLICATION NUMBER: US/10/400,708
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-400-708-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 91

RESULT 7
US-10-153-207-6
; Sequence 6, Application US/10153207
; Publication No. US2003015303A1
; GENERAL INFORMATION:
; APPLICANT: James A. Wells
; APPLICANT: Brian C. Cunningham
; TITLE OF INVENTION: GROWTH HORMONE VARIANTS
; FILE REFERENCE: 669.12-US-C7
; CURRENT APPLICATION NUMBER: US/10/153,207
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: 08/479,884
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/190,723
; PRIOR FILING DATE: 1994-02-02
; PRIOR APPLICATION NUMBER: 07/960,227
; PRIOR FILING DATE: 1992-10-13
; PRIOR APPLICATION NUMBER: 07/875,204
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/428,066
; PRIOR FILING DATE: 1989-10-26
; PRIOR APPLICATION NUMBER: 07/264,611
; PRIOR FILING DATE: 1988-10-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-207-6

Query Match          97.9%   Score 460; DB 12; Length 214;
Best Local Similarity 98.9%   Pred. No. 9e-45;
Matches 90: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 24 FFTPISRLFDNAMLRAHRLHQLAFDTYQFEAYIPKEOKYSFLQNPTSLSFSES IPT 83

QY 62 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 84 PSNREETQOKSNLELLKISLLLIQSWLEPVQ 114

RESULT 8
US-09-929-918-9
```


QY 2 EPTPLSLIEDNMLRARRHQLAEDTYQRFEAYIKRKVKYKFLONGTSLSSSESYIT 61
:
DB 27 EDRTDLSPIFNASIPADIRFOTAEITYQEGDNVHWPKVYKFLONGTSLSSSESYIT 69

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/087,678
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: GERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132

```

: REFERENCE/DOCKET NUMBER: PR-0555 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (850) 853-0555
: TELEFAX: (850) 843-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 217 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: c406987
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-850-887.3

Query Match      84.3%, Score 396, DB 0, Length 217:
Best Local Similarity 84.6%, Prod. No. 1.9e-37:
Matches 77: Conservative 6; Mismatches 8; Indels 0; Gaps 0:

QY      2 FPIPLSKLFDNMIKRAHRLHQLAFETYQPFEEAYPKKQKYSFLQNTQTSISFSSESTPT 41
Db      27 FPIPLSWLENTAVFRAHHLKLAFTYPKLEAYPKKQKYSFLRNPQTSICFSSESTPT 86

QY      62 PSNKEETQCKSNLEKRIKSLKQSWLEPVQ 92
Db      87 PSNKEETQCKSNLELHISLQSWLEPVQ 117

```

Search completed: September 15, 2003, 12:23:15
Job time : 39.5806 secs

A:Residues: 27-94:96-217 <LNC>
R:Niall, H.D.
Nature New Biol. 230, 90-91, 1971
A:Title: Revised primary structure for human growth hormone.
A:Reference number: A93397; MUID:71139765; PMID:5279045
A:Accession: A93397
A:Molecule type: protein
A:Residues: 27-51 <NIA>
R:Niall, H.D.; Honan, M.L.; Sayer, R.; Rosenblum, I.Y.; Greenwood, F.C.
Proc. Natl. Acad. Sci. U.S.A. 63, 866-869, 1971
A:Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution
A:Reference number: A93778; MUID:71153968; PMID:5279528
A:Accession: A93778
A:Molecule type: protein
A:Residues: 119-120:157-159 <NIA>
R:Niall, H.D.
in Prolactin and Carcinogenesis, Proc. Fourth Toronto Workshop Prolactin, Griffiths, K.A.
A:Title: The chemistry of the human lactogenic hormones.
A:Reference number: A94427
A:Contents: annotation: somatotropin revision
R:Bewley, T.A.; Dixon, J.S.; Li, C.H.
Int. J. Pept. Protein Res. 4, 231-287, 1972
A:Title: Sequence comparison of human pituitary growth hormone, human chorionic somatomedin
A:Reference number: A91764; MUID:7392028; PMID:4655454
A:Accession: A91764
A:Molecule type: protein
A:Residues: 27-217 <BEW>
R:Lewis, G.J.; Borowald, L.F.; Lewis, G.J.
Biochem. Biophys. Res. Commun. 92, 511-516, 1983
A:Title: The 20,000-dalton variant of human growth hormone: location of the amine acid
A:Reference number: A90217; MUID:60130196; PMID:73356475
A:Contents: somatotropin, 20K short variant
A:Accession: A90217
A:Molecule type: protein
R:Chapman, G.E.; Rogers, K.M.; Brittain, T.; Broadshaw, R.A.; Bates, M.J.; Primer, G.; G.
J. Biol. Chem. 256, 2395-2401, 1981
A:Title: The 20,000 molecular weight variant of human growth hormone: Preparation and se
A:Reference number: A92311; MUID:61117361; PMID:7462247
A:Contents: somatotropin, 20K short variant
A:Accession: A92311
A:Molecule type: protein
R:Singh, R.N.P.; Scavey, B.K.; Lewis, G.J.; Lewis, G.J.
J. Protein Chem. 2, 425-436, 1983
A:Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
A:Reference number: A61466
A:Accession: A61466
A:Molecule type: protein
A:Residues: 27-63 <SIN>
R:Robson, V.M.C.; Rae, I.D.; NG, F.
Biol. Chem. Hoppe-Seyler 371, 423-431, 1990
A:Title: Identification of the aspartamide structure in a previously-reported peptide.
A:Reference number: S09685; MUID:90334745; PMID:2338679
A:Accession: S09585
A:Molecule type: protein
A:Residues: 27-34:57-36-47 <ROB>
R:de Vos, A.M.; Ullsch, M.; Kossiakoff, A.A.
Science 255, 306-312, 1992
A:Title: Human growth hormone and extracellular domain of its receptor: crystal structure
A:Reference number: A41728; MUID:92196577; PMID:1549776
A:Contents: annotation: X-ray crystallography, 2.8 angstroms
R:Gray, G.L.; Balbridge, J.S.; McKown, K.S.; Heyneker, H.L.; Chang, C.N.
Gene 39, 247-254, 1985
A:Title: Periplasmic production of correctly processed human growth hormone in Escherich
A:Reference number: 141126; MUID:86137393; PMID:3912251
A:Accession: 184543
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-26 <RES>
A:Cross-references: GB:M44360; MUI:9163150; PIDN:AAA52554.1; PID:9163150

C:Comment: The gene for this hormone is transcribed only in somatotrophic cells of
C:Comment: About 90% of somatotropin is the 22K long form.
C:Genetics:
A:Gene: GHB:GHI
A:Cross-references: GDB:119982; OMIM:139250
A:Map position: 17q23.1-17q23.3
A:Introns: 4/1; 57/3; 97/3; 152/3
C:Superfamily: prolactin
C:Keywords: alternative splicing; hormone; pituitary
F:1-26/UCrain: signal sequence #status experimental <SIG>
F:27-217/Product: somatotropin 1, long form #status experimental <SOL>
F:27-69/Product: growth hormone 5K peptide #status experimental <SKP>
F:27-57,73-217/Product: somatotropin 1, short form #status experimental <SOS>
F:79-191,208-215/Disulfide bonds: #status experimental

Query Match 97.9%; Score 460; DB 1; Length 217;
Best Local Similarity 98.9%; Pred. No. 4.4e-42;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRLLFNAMLAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPTSLSFSESPT 61
|||||
DB 27 FPTPLSRLLFNAMLAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPTSLSFSESPT 86
|||||

QY 62 PSNRETOQSKNLELRISLLLIQSWLEPVQ 92
|||||
DB 87 PSNRETOQSKNLELRISLLLIQSWLEPVQ 117
|||||

RESULT 2
157410
somatotropin - rhesus macaque
N:Alternate names: growth hormone
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: 167410; A05094
R:Goebel, T.G.; Burnip, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related compl
A:Reference number: 153267; MUID:94008724; PMID:8404617
A:Accession: 167410
A:Status: translated from GH/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GRLJ8556; MUI:9293114; PIDN:AAA18842.1; PID:9293115
R:Li, C.H.; Chung, D.; Lam, H.W.; Stein, S.
Arch. Biochem. Biophys. 245, 287-291, 1986
A:Title: The primary structure of monkey pituitary growth hormone.
A:Reference number: A05094; MUID:86229460; PMID:3080959
A:Accession: A05094
A:Molecule type: protein
A:Residues: 27-99:101-178:180-217 <LIC>
A:Note: the monkey species is not identified in the reference
R:Raben, M.S.
Science 125, 883-884, 1957
A:Title: Preparation of growth hormone from pituitaries of man and monkey.
A:Reference number: A44774
A:Contents: annotation: identification of source organism
C:Superfamily: prolactin

Query Match 97.9%; Score 460; DB 2; Length 217;
Best Local Similarity 98.9%; Pred. No. 4.4e-42;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRLLFNAMLAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPTSLSFSESPT 61
|||||
DB 27 FPTPLSRLLFNAMLAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPTSLSFSESPT 86
|||||

QY 62 PSNRETOQSKNLELRISLLLIQSWLEPVQ 92
|||||
DB 87 PSNRETOQSKNLELRISLLLIQSWLEPVQ 117
|||||

RESULT 3

```

STCUTV
somatotropin 2 precursor - human
N:Alternate names: growth hormone variant; hGH-V; placental somatotropin
N:Contains: somatotropin 2, long splice form; somatotropin 2, short splice form
C:Species: Homo sapiens (man)
C:Date: 17-Dec-1982 #sequence_revision 10-Feb-1995 #text_change 21-Jul 2000
C:Accession: D32435; B28072; A01511; I52104; A60711
C:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gallinas, R.E.; Seeburg, P.H.
Genomics 4, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, ecology, and evolution
A:Reference number: A32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: D32435
A:Molecule type: cDNA
A:Residues: 1-217 <CDS>
A:Cross-references: GB:103071; NID:183346; PID:AAA52521; PID:183346
R:Coake, N.E.; Ray, J.; Emery, J.G.; Liebhaber, S.A.
J. Biol. Chem. 263, 9001-9006, 1988
A:Title: Two distinct species of human growth hormone variant mRNA in the human placenta
A:Reference number: A92725; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: B28072
A:Molecule type: mRNA
A:Residues: 1-217 <CDS>
R:Seeburg, P.H.
DNA 1, 239-249, 1982
A:Title: The human growth hormone gene family: nucleotide sequences show recent divergence
A:Reference number: A01511; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: A01511
A:Molecule type: cDNA
A:Residues: 1-34, 136-217 <SP>
R:Ignot, A.; Scippo, M.L.; Franke, F.; Heinen, G.
Arch. Int. Physiol. Biochim. 96, 53-67, 1988
A:Title: Cloning and nucleotide sequence of placental hGH-V cDNA
A:Reference number: I52104; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: I52104
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <IG>
A:Cross-references: GB:M38451; NID:183379; PID:AAA35891; PID:183379
R:Frankenne, F.; Scippo, M.L.; Van Beunen, J.; Ignot, A.; Heinen, G.
J. Clin. Endocrinol. Metab. 71, 15-19, 1990
A:Title: Identification of placental human growth hormone as the growth hormone-V gene
A:Reference number: A60711; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: A60711
A:Molecule type: protein
A:Residues: 27-44; 46-57 <FR>
A:Experimental source: tissue placenta
A:Note: partial glycosylation was demonstrated by lectin binding
C:Comment: This gene is expressed by the placenta.
C:Genetics:
A:Gene: GDB:GH2
A:Cross-references: GDB:119983; OMIM:139240
A:Map position: 17q22-17q24
A:Introns: 4/1; 57/3; 97/3; 152/3
C:Superfamily: prolactin
C:Keywords: alternative splicing; glycoprotein; hormone; placenta
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-57.73-217/Product: somatotropin 2, long splice form #status predicted <SIG>
F:79-191,208-215/Dissulfide bonds: #status predicted
F:156/Binding site: carbohydrate (Asn) (covariant) #status predicted

Query Match 89.8% Score 422; DB 1; Length 217;
Best Local Similarity 92.3% Pred. No. 5,50-38;
Matches 84; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 41
DB 27 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 46
QY 62 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 92
DB 97 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 117

```

```

RESULT 4
STCUTV
somatotropin 2 precursor, splice form 2 - human
N:Alternate names: growth hormone variant-2; placental somatotropin form 2
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1989 #sequence_revision 10-Feb-1995 #text_change 02-Sep-1997
C:Accession: A28072
R:Coake, N.E.; Ray, J.; Emery, J.G.; Liebhaber, S.A.
J. Biol. Chem. 263, 9001-9006, 1988
A:Title: Two distinct species of human growth hormone variant mRNA in the human placenta
A:Reference number: A92725; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: A28072
A:Molecule type: mRNA
A:Residues: 1-255 <CDS>
A:Note: an alternative splice junction for intron 4 is used
C:Genetics:
A:Gene: GDB:GH2
A:Cross-references: GDB:119983; OMIM:139240
A:Map position: 17q22-17q24
A:Introns: 4/1; 57/3; 97/3; 152/3
C:Superfamily: prolactin
C:Keywords: alternative splicing; hormone; placenta
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-255/Product: somatotropin 2 splice form 2 #status predicted <NAT>

Query Match 89.8% Score 422; DB 1; Length 256;
Best Local Similarity 92.3% Pred. No. 6,7e-36;
Matches 84; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 61
DB 27 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 86
QY 62 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 92
DB 87 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 117

RESULT 5
somatotropin - rhesus macaque
N:Alternate names: growth hormone
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: 167411
R:Golos, T.G.; Burnington, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complex
A:Reference number: 153257; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: 167411
A:Status: preliminary; translated from GB/EMBL/DDHJ
A:Molecule type: mRNA
A:Residues: 1-217 <Res>
A:Cross-references: GB:116555; NID:G253116; PID:AAA20180.1; PID:G293117
C:Superfamily: prolactin

Query Match 85.5% Score 402; DB 2; Length 217;
Best Local Similarity 85.7% Pred. No. 7,8e-36;
Matches 78; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 61
DB 27 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 86
QY 62 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 92
DB 87 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 117

RESULT 6
somatotropin - rhesus macaque
N:Alternate names: growth hormone
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: 167411
R:Golos, T.G.; Burnington, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complex
A:Reference number: 153257; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435; M32435
A:Accession: 167411
A:Status: preliminary; translated from GB/EMBL/DDHJ
A:Molecule type: mRNA
A:Residues: 1-217 <Res>
A:Cross-references: GB:116555; NID:G253116; PID:AAA20180.1; PID:G293117
C:Superfamily: prolactin

Query Match 85.5% Score 402; DB 2; Length 217;
Best Local Similarity 85.7% Pred. No. 7,8e-36;
Matches 78; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 61
DB 27 FPTPLSLFNDNMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLONPQTSLSFSSEIPT 86
QY 62 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 92
DB 87 PSNRKSTOCKSNLELLRLSILLIQSWLEPVQ 117

```

```

C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision: 31-May-1996 #text_change 16-Jul-1999
C:Accession: I67409
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A:Reference number: 153267; MUID:94008724; PMID:8404617
A:Accession: I67409
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:I16554; NID:9293112; PIDN:AAA:8341.1; PID:9293113
C:Superfamily: prolactin

Query Match      84.3%; Score 397; DB 2; Length 217;
Best Local Similarity 83.3%; Pred. No. 3.4e-35;
Matches 75; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 3 PTIPLSRLFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 28 PSVPLSLRFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 67
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 SNRETOQKSNLELRISLLIQSWLEPVQ 92
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 68 SNRETOQKSNLELRISLLIQSWLEPVQ 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
I67408
chorionic somatomammotropin-2 - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision: 31-May-1996 #text_change 16-Jul-1999
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A:Reference number: 153267; MUID:94008724; PMID:8404617
A:Accession: I67408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <RES>
A:Cross-references: GB:I16553; NID:9293110; PIDN:AAA:18849.1; PID:9293111
C:Superfamily: prolactin

Query Match      84.3%; Score 396; DB 2; Length 212;
Best Local Similarity 82.2%; Pred. No. 3.4e-35;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PTIPLSRLFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 23 PSVPLSLRFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 SNRETOQKSNLELRISLLIQSWLEPVQ 92
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 83 SNRETOQKSNLELRISLLIQSWLEPVQ 112
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
I53267
chorionic somatomammotropin-1 - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-May-1996 #sequence_revision: 31-May-1996 #text_change 16-Jul-1999
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A:Reference number: 153267; MUID:94008724; PMID:8404617
A:Accession: I53267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:I16552; NID:9293108; PIDN:AAA:16839.1; PID:9293109
C:Superfamily: prolactin

```

```

Query Match      84.3%; Score 396; DB 2; Length 217;
Best Local Similarity 82.2%; Pred. No. 3.4e-35;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PTIPLSRLFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 62
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 28 PSVPLSLRFDNMLRAHRLHQAFDTYQFEFAYIPKCKYSLFQNPQTSLSFSSSITPT 67
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 SNRETOQKSNLELRISLLIQSWLEPVQ 92
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 86 SNRETOQKSNLELRISLLIQSWLEPVQ 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
I53268
choriomammotropin A precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 08-Dec-2000
C:Accession: G32435; A94422; A93833; A93192; A90054; A94427; A61283; 155229
R:Chen, P.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saidana, H.A.; Gelinas, R.E.; Seebur
g, G. 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolutio
A:Reference number: A32435; MUID:89307277; PMID:2744760
A:Accession: G32435
A:Molecule type: DNA
A:Residues: 1-217 <CHE>
A:Cross-references: GB:J03071; NID:9183148; PIDN:AAA:52551.1; PID:9183151
R:Goodman, H.M.; Denoto, F.; Fiddes, J.C.; Hallowell, R.A.; Page, G.S.; Smith, S.;
in Mobilization and Reassembly of Genetic Information, Scott, W.A.; Werner, R.; Jos
A:Reference number: A94422
A:Accession: A94422
A:Molecule type: mRNA
A:Residues: 1-217 <GOO>
R:Tanaka, M.; Masuda, K.; Watabiki, M.; Yamakawa, M.; Shimizu, K.; Nagai, J.; Nakas
higuchi, H. 16, 287-292, 1988
A:Title: cDNA cloning of human chorionic somatomammotropin-1 mRNA whose transcripti
A:Reference number: 152342; MUID:88209096; PMID:2835050
A:Accession: I52342
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3 <IAN>
A:Cross-references: GB:M35413; NID:9506622
R:Sherwood, E.M.; Bernstein, Y.; Schecter, I.
Proc. Natl. Acad. Sci. U.S.A. 76, 3819-3823, 1979
A:Title: Primary structure of the NH-2-terminal extra piece of the precursor to hum
A:Reference number: A3833; MUID:83034970; PMID:291043
A:Accession: A3833
A:Molecule type: protein
A:Residues: 1-3-26 <SH>
A:Experimental source: placenta
R:Shine, J.; Seeburg, P.H.; Martial, J.A.; Baxter, J.D.; Goodman, H.M.
Nature 270, 494-499, 1977
A:Title: Construction and analysis of recombinant DNA for human chorionic somatomam
A:Reference number: A93192; MUID:75071761; PMID:593369
A:Accession: A93192
A:Molecule type: DNA
A:Residues: 50-217 <SH>
A:Experimental source: placenta
R:Li, C.H.; Dixon, J.S.; Chung, D.
Arch. Biochem. Biophys. 155, 95-110, 1973
A:Title: Amino acid sequence of human chorionic somatomammotropin.
A:Reference number: A90054; MUID:73201971; PMID:4712450
A:Accession: A90054
A:Molecule type: protein
A:Residues: 27-217 <LIC>
A:Experimental source: placenta
R:Niall, R.D.
in Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop Prolactin, Griffiths
A:Title: The chemistry of the human lactogenic hormones.
A:Reference number: A94427
A:Accession: A94427

```

N: Alternate names: chorionic somatomammotropin 2
C: Species: Homo sapiens (man)
C: Date: 29-Dec-1989 #sequence_revision 29-Dec-1989 #text_change 16-Jul-1999
C: Accession: E32435
R: Chen, E.Y.; Liaw, Y.C.; Smith, D.H.; Herrera-Saldana, H.A.; Gelinas, R.E.; Seeburgenomics 4; 479-497, 1989
A: Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A: Reference number: A32435; MUID:89307277; PMID:2744760
A: Accession: E32435
A: Status: preliminary
A: Category: protein
A: Molecule type: DNA
A: Residues: 1-217 <HE>
A: Cross-references: GR:JC3071; NID:g183148; PIDN:AAA52553.1; PDC:g183153
C: Genes:
A: Gene: GHR:GSH2
A: Cross-references: GDB:l19833; OMIM:l18820
A: Map position: 17q22-q24
C: Superfamily: prolactin

Query Match 81.1% Score 381; DB 2; Length 217;
Best Local Similarity 82.0%; Pred. No. 1,4e-33;
Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TPLSLFNFAMRAHRLHQAFDYQEERAYIPKEKYSEFLNQPTLSLSESIIPTS 63
DE :|||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||
29 TVPLSRLEFHAMLOAHRAHQAIAITDYEETIPIKDKYSFLHDSQSIFSDSIPTS 88
QY 64 NNEETOCKSNLELRISLLIQSWLEPVQ 92
DE :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
89 NNEETOCKSNLELRISLLIQSWLEPVR 117

RESULT 11
A26449
Choriomamotropin precursor (allele hCS-3) - human
C: Species: Homo sapiens (man)
C: Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 28-Jul-1995
C: Accession: A26449
R: Birb, H.; Kimmelman, J.; Birnbaum, M.J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.; DNA 6; 59-70, 1987
A: Title: The human growth hormone gene locus: structure, evolution, and allelic variation
A: Reference number: A26449; MUID:87161235; PMID:3030680
A: Accession: A26449
A: Molecule type: DNA
A: Residues: 1-215 <HIR>
C: Superfamily: prolactin
E: 1-26/Domain: signal sequence *status predicted <SIG>
E: 27-215/Product: choriomamotropin, hCS-3 allele *status predicted <MAT>

Query Match 76.5% Score 359.5; DB 2; Length 215;
Best Local Similarity 80.5%; Pred. No. 2.9e-31;
Matches 70; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 4 TPLSLFNFAMRAHRLHQAFDYQEERAYIPKEKYSEFLNQPTLSLSESIIPTS 63
DE :|||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||
29 TVPLSRLEFHAMLOAHRAHQAIAITDYEETIPIKDKYSFLHDSQSIFSDSIPTS 88
QY 64 NNEETOCKSNLELRISLLIQSWLEPVQ 90
DE :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
89 NNEETOCKSNLELRISLLIQSWLEPVR 114

RESULT 12
A49159
Somatotropin - golden hamster
N: Alternate names: growth hormone
C: Species: Mesocricetus auratus (golden hamster)
C: Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C: Accession: B49159
R: Southard, J.N.; Sanchez-Jimenez, F.; Campbell, G.F.; Talamantes, F.; Endocrinology 129; 2965-2971, 1991
A: Title: Sequence and expression of hamster prolactin and growth hormone messenger RNA
A: Reference number: A49159; MUID:92063850; PMID:1954881

GenInfo version 5.1.1
Copyright (c) 1993 - 2003 CompuLink Ltd.

EM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:06 : Search time 0:17:48 seconds
(without alignment)
504,637 BLASTN Cell updates/sec

Title: US-09-423-100-2

Perfect score: 470

Sequence: 1 MFPIPLSKFENAMRAHP.....NFTLRSLNLSLQNMDEWQ 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026765 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 6

Maximum DB seq length: 200000000

Post-processing: Minimum Match 60

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	460	97.9	217	1	SOMA_HUMAN
2	460	97.9	217	1	SOMA_MACMU
3	460	97.9	217	1	SOMA_PANTR
4	433	92.3	217	1	SOMA_SALIB
5	432	91.9	217	1	SOMA_GALLA
6	430	91.5	217	1	SOMA_PANTR
7	422	89.8	217	1	SOMA_HUMAN
8	396	84.3	217	1	SOMA_MACMU
9	391	81.1	217	1	PEL_HUMAN
10	310.5	66.1	215	1	SOMA_MESAU
11	307.5	65.4	190	1	SOMA_BALBO
12	304.5	64.8	216	1	SOMA_MOUSE
13	302.5	64.4	216	1	SOMA_HORSE
14	302.5	64.4	216	1	SOMA_RABIT
15	302.5	64.4	216	1	SOMA_RAT
16	302.5	64.4	217	1	SOMA_GALSE
17	302.5	64.4	217	1	SOMA_NYCPV
18	301.5	64.1	190	1	SOMA_GUXAF
19	301.5	64.1	216	1	SOMA_PANPA
20	301.5	64.1	216	1	SOMA_FELCA
21	301.5	64.1	216	1	SOMA_PYG
22	299.5	63.7	215	1	SOMA_MOUSE
23	297.5	63.3	190	1	SOMA_LAKKA
24	295.5	62.9	190	1	SOMA_VULVU
25	291.5	62.0	215	1	SOMA_MONIC
26	291.5	62.0	215	1	SOMA_TRIUV
27	289.5	61.6	217	1	SOMA_BOVIN
28	289.5	61.6	217	1	SOMA_CSEEL
29	289.5	61.6	217	1	SOMA_SHEEP
30	282.5	60.1	217	1	SOMA_RABBO
31	278.5	59.3	215	1	SOMA_MELGA
32	275.5	58.6	216	1	SOMA_CHICK
33	274.5	58.4	217	1	SOMA_STRCA

ALIGNMENTS

RESULT 1

SOMA_HUMAN
 ID: SOMA_HUMAN STANDARD: PRT: 217 AA.
 AC: F01241; 014405; Q16631; Q9H521; Q9UMJ7; Q9UNC5;
 DT: 21-JUL-1986 (Rel. 01. Created)
 DI: 01-MAR-1992 (Rel. 21. Last sequence update)
 UT: 15-SEP-2003 (Rel. 42. Last annotation update)
 DE: Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
 GN: SH1.
 OS: Homo sapiens (Human).
 OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX: NCBI_TaxID=9606;
 RN: 1.
 RP: SEQUENCE FROM N.A. (ISOFORM 1).
 RA: MEDLINE-80034477; PubMed-386281;
 RT: "Roskam W., Rougeon F.;
 RI: "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."
 RJ: Hormone structural gene."
 RL: Nucleic Acids Res. 7:305-320(1979).
 RN: 12.
 RP: SEQUENCE FROM N.A. (ISOFORM 1).
 RA: MEDLINE-79203293; PubMed-377496;
 RT: "Martini J.A., Hallewell R.A., Baxter J.D., Goodman H.M.;
 RI: "Human growth hormone: complementary DNA cloning and expression in bacteria."
 RJ: Science 205:602-607(1979).
 RN: 13.
 RP: SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.
 RA: MEDLINE-82014939; PubMed-6269091;
 RT: "Deroto F.M., Moore D.D., Goodman H.M.;
 RI: "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."
 RJ: Nucleic Acids Res. 9:3719-3730(1981).
 RN: 14.
 RP: SEQUENCE FROM N.A.
 RA: MEDLINE-83182010; PubMed-7169009;
 RT: "Seeburg P.H.;
 RI: "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."
 RJ: DNA 1:239-249(1982).
 RN: 15.
 RP: SEQUENCE FROM N.A.
 RA: MEDLINE-89407277; PubMed-2744760;
 RT: "Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.,
 RI: "Gelinias R.B., Seeburg P.H.;
 RJ: "The human growth hormone locus: nucleotide sequence, biology, and evolution."
 RJ: Genomics 4:479-487(1989).
 RN: 16.
 RP: SEQUENCE FROM N.A. (ISOFORM 3).
 RA: TISSUE-Pituitary;
 RT: "Gu G., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;
 RI: "A novel gene expressed in human pituitary."
 RJ: Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

P5755 crocodylus
 P34005 ctenomys
 P11228 anas platyr
 P26773 acipenser g
 P26774 acipenser g
 P79885 lepisosteus
 P10813 xanopus lae
 P10813 xanopus lae
 Q73849 bufo marinus
 P34006 priodonace gl
 P2856 xanopus lae
 Q73848 propterus

34 272.5 58.0 190 1 SOMA_CRONO
 35 268.5 57.1 191 1 SOMA_CHEMY
 36 261 55.5 216 1 SOMA_ANAPI
 37 257.5 54.8 190 1 SOM1_ACIGU
 38 257.5 54.8 190 1 SOM2_ACIGU
 39 247.5 52.7 215 1 SOMA_LEPOS
 40 238.5 51.0 214 1 SOMA_XENIA
 41 238.5 50.7 215 1 SOMA_XENIA
 42 225.5 48.2 215 1 SOMA_BUPMA
 43 225.5 48.0 183 1 SOMA_PRIGL
 44 219.5 46.7 208 1 SOMB_XENIA
 45 218.5 45.5 206 1 SOMA_PROAN

RN [7] SEQUENCE FROM N.A. (ISOFORM 4).
 RP TISSUE-Pituitary.
 RX MEDLINE=20402571; PubMed=19911945;
 RA Hu R.-M., Han Z.-G., Song H.-D., Huang Q.-B., Ren S.-X.,
 GU Y.-S., Huang C.-H., Li Y.-S., Jiang C.-L., Fu G., Zhang T.-B.,
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong K., Ye M., Zhou C.-H.,
 RA Xu S.-H., Gu J.-X., Shi J.-X., Jin W.-K., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.,
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 axis and full-length cDNA cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN [8]
 RP REVIEW.
 RX MEDLINE=99321812; PubMed=10353484;
 RA Baumann G.;
 RI "Growth hormone heterogeneity in human pituitary and plasma.";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RN [19]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=88190073; PubMed=3447173;
 RA Cohen P.F., Kuntz L.P.;
 RI "Prediction of the three-dimensional structure of human growth
 hormone.";
 RL Proteins 2:162-165(1987).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=92195577; PubMed=1549776;
 RA de Vos A.M., Ullsch M., Kossiakoff A.A.;
 RI "Human growth hormone and extracellular domain of its receptor:
 crystal structure of the complex.";
 RL Science 255:305-312(1992).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=95075452; PubMed=7584244;
 RA Somers W., Ullsch M., de Vos A.M., Kossiakoff A.A.;
 RI "The X-ray structure of a growth hormone-prolactin receptor complex.";
 RL Nature 372:478-481(1994).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RA Chantalat L., Chirgadze N.Y., Jones N., Korber F., Navaza J.,
 Pavlovsk A.G., Wlodawer A.;
 RI "The crystal-structure of wild-type growth-hormone at 2.5-A
 resolution.";
 RL Protein Pept. Lett. 2:333-340(1995).
 RN [23]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97133023; PubMed=6943276;
 RA Sundstroem M., Lundqvist T., Roedin J., Glebel L.B., Milligan D.,
 Norstedt G.;
 RI "Crystal structure of an antagonist mutant of human growth hormone,
 g120R, in complex with its receptor at 2.9-A resolution.";
 RL J. Biol. Chem. 271:32197-32203(1996).
 RN [24]
 RP VARIANT CYS-105.
 RX MEDLINE=99318053; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RA "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [25]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role in stimulating body growth is to stimulate the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates

[illegible]

RT "Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional promoter elements." J. Biol. Chem. 259:13131-13138(1984).

RN [2]

RP SEQUENCE FROM N.A. (GENE CSH2).

RX MEDLINE=87161235; PubMed=3030580;

RA Hirt H., Kime-man J., Birnbaum M.J., Chen E.Y., Seeburg P.H., Eberhardt N.E., Barja A.

RT "The human growth hormone gene locus: structure, evolution, and allelic variations." DNA 6:59-70(1987).

RL [3]

RN SEQUENCE FROM N.A.

RP MEDLINE=83163916; PubMed=6300056;

RA Barrera-Saldana H.A., Secor P.H., Saunders G.F.

RT "Two structurally different genes produce the same secreted human placental lactogen hormone." J. Biol. Chem. 259:3787-3793(1984).

RL [4]

RP SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).

RX MEDLINE=83307277; PubMed=2744769;

RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinis E.E., Seeburg P.H.

RT "The human growth hormone locus: nucleotide sequence, biology, and evolution." Genomics 4:479-497(1989).

RL [5]

RP SEQUENCE.

RX MEDLINE=83182610; PubMed=7169009;

RA Seeburg P.H.

RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone." DNA 1:239-243(1982).

RL [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta, and Uterus;

RX MEDLINE=22386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.S., Klausner R.D., Collins F.S., Wagner L., Shenmen C.Y., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Datchenko L., Karusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.H., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson P.D., McLachy S., Bosak S.A., McEwan P.J., McKernan A.J., Malek J.A., Gunaratne P., Richards S., Worley K.C., Hale S., Garcia A.M., Gay D., Huiyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner N., Madan A., Rochevsky S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Battersfield V.S.N., Kravitzski M.J., Skalska E., Smolins D.E., Schnerch A., Schein J.E., Jones S.J.M., Mitra N.A.

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).

RL [7]

RP SEQUENCE OF 50-217 FROM N.A.

RX MEDLINE=78071761; PubMed=553368;

RA Shiao J., Seeburg P.H., Martial J.A., Baxter J.D., Goodman R.E.

RT "Construction and analysis of recombinant DNA for human chorionic somatomammotropin." Nature 270:494-499(1977).

RL [8]

RP SEQUENCE OF 27-217.

RX MEDLINE=73201971; PubMed=4712450;

RA Li C.H., Dixon J.S., Chang D.

RT "Amino acid sequence of human chorionic somatomammotropin." Arch. Biochem. Biophys. 155:95-110(1973).

RL [9]

RP SEQUENCE OF 27-117.

RX MEDLINE=72016113; PubMed=5286363;

RA Sherwood L.M., Handwerger S., McLaurin W.D., Lanner M.; "Amino-acid sequence of human placental lactogen." Nature New Biol. 233:59-61(1971).

RN [10]

RP ERRATUM.

RA Sherwood L.M., Handwerger S., McLaurin W.D., Lanner M.; Nature New Biol. 235:64-64(1972).

RN [11]

RP INTERCHAIN DISULFIDE BONDS.

RX MEDLINE=79173081; PubMed=438159;

RA Schneider A.B., Kowalski A., Russell J., Sherwood L.M.; "Identification of the interchain disulfide bonds of dimeric human placental lactogen." J. Biol. Chem. 254:3782-3787(1979).

RL [12]

RP FUNCTION: SIMILAR TO THAT OF SOMATOTROPIN.

CC [13] SUBCELLULAR LOCATION: Secreted.

CC [14] MISCELLANEOUS: THE SEQUENCE OF CSH1 IS SHOWN.

CC [15] SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: V00573; CAA23836.1; -

DR EMBL: J06289; AAA98747.1; -

DR EMBL: K02401; AAA52115.1; -

DR EMBL: M15894; AAA52116.1; -

DR EMBL: J03071; AAA52551.1; -

DR EMBL: J06118; AAA94621.1; -

DR EMBL: BC002717; AA822717.1; -

DR EMBL: BC005921; AA835921.1; -

DR EMBL: BC020756; AA820756.1; -

DR PIR: A26449; A26449.

DR PIR: C32435; LCHUC.

DR RSP: P01241; IA22.

DR Genew: HGNC:2440; CSH1.

DR Genew: HGNC:2441; CSH2.

DR MIN: 150200; -

DR GO: GO:0007365; P:pregnancy; TAS.

DR InterPro: IPR001400; Somatotropin.

DR Pfam: PF00103; hormone; 1

DR PROSITE: PS00256; SOMATOTROPIN_1; 1.

DR PROSITE: PS00338; SOMATOTROPIN_2; 1.

DR Hormone: Placenta; Multigene family; Signal.

KW SIGNAL 1 26

FT CHAIN 27 217 LACTOGEN.

FT DISULFID 79 191

FT DISULFID 208 215

FT DISULFID 208 208

FT DISULFID 215 215

FT VARIANT 3 3

FT VARIANT 104 105

FT CONFLICT 84 84 1 -> T (IN REF. 9).

FT CONFLICT 95 95 MISSING (IN REF. 9).

FT CONFLICT 126 116 MISSING (IN REF. 9).

FT CONFLICT 134 136 SDD -> BBS (IN REF. 9).

FT SEQUENCE 217 AA; 25920 MW; 235B0DC7A713F431 CRC64;

Query Match 81.1% Score 381; DB 1; Length 217;

Best local similarity 62.0% Pred. No. 1.4e-42;

Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TPLSRLFDNMLRAHRLHOLAIDTQVEFEAVIPKEQKYSFLQNPOITS-SFSESPTPS 63

DB 29 TVPLSRLFDHMLQAHRAHOLAIDTQVEFEVTPKDKKYSFLQDSQITSFSDSPTPS 88

QY 64 NREFTQOKSNLELLR-SLLLIQSWLEPVQ 92


```

RL J. Biol. Chem. 260:9574-9579(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN:F2IDU; TISSUE=liver;
RX MEDLINE-9519489; PubMed=8547448;
RA Das P., Meyer L., Seyfert J.N., Brockmann G., Schwertin M.
RT "Structure of the growth hormone-encoding gene and its promoter in
   mice."
RL Gene 169:209-213(1996).
CC -!- FUNCTION: Plays an important role in growth control. Its major
   role in stimulating body growth is to stimulate the liver and
   other tissues to secrete IGF-1. It stimulates both the
   differentiation and proliferation of myoblasts. It also stimulates
   amino acid uptake and protein synthesis in muscle and other
   tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation
   at the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announcement/
   or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X02991; CAA26650.1; .
DR FMBL: 245653; CAA86653.1; .
DR PIR: B23911; STMS.
DR HSSP: P01246; IBS7.
DR MGD: MGI:95707; Ch.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA; 24716 MW; 98656A3AE25D65FC CRC64;
Query Match 64.8%; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 2e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;
QY 2 FPTPLSRFDNAMLRAHRLAQAFDYQFEERAYIPKEQKYSFLONPQTSLSFSSESPT 61
DB 27 FPAPPLSLSFANVLRADQLHQLAADTYKEFERAYIPKEQKYS-IQNAQAFCSETIPA 55
QY 62 PSNREETQOKSNLELRISLLLIQSWLEPVQ 92
DB 86 PTGKEAQQRDMELLRFSLIIQSWLGPVQ 116
RESULT 13
SOMA_HORSE
ID SOMA_HORSE STANDARD; PRT: 216 AA.
AC P01245.
DI 01-JUL-1986 (Rel. 21, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutrachia; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE-9426171; PubMed=8206392;
RA Ascacio-Martinez J.A., Barrera-Saldana H.A.;

```

```

*Sequence of a cDNA encoding horse growth hormone.*;
Gene: 143:299-300(1994).
RN [2]
RP SEQUENCE OF 27-216.
RX MEDLINE-77005410; PubMed:965151;
RA Zakin M.M., Poskus E., Langton A.A., Ferrara P., Santone J.A.,
RA Dellacha J.M., Paladini A.C.;
RT "Primary structure of equine growth hormone.";
RL Ist. C. Pept. Protein Res. 8:435-444(1976).
RN [3]
RP PRELIMINARY SEQUENCE OF 27-216.
RX MEDLINE-74020362; PubMed:4747843;
RA Zakin M.M., Poskus E., Dellacha J.M., Paladini A.C., Santone J.A.;
RT "The amino acid sequence of equine growth hormone.";
RL FEBS Lett. 34:353-355(1973).
RN [4]
RP SEQUENCE OF 68-95 AND 183-216.
RA Zakin M.M., Poskus E., Dellacha J.M., Paladini A.C., Santone J.A.;
RT "Amino acid sequences around the cysteine residues in equine growth
   hormone.";
RL FEBS Lett. 25:77-82(1972).
RN [5]
RP SEQUENCE OF 202-216.
RX MEDLINE-68368390; PubMed:4876100;
RA Gliver L., Hartree A.S.;
RT "Amino acid sequences around the cysteine residues in horse growth
   hormone.";
RL Biochem. J. 109:19-24(1968).
CC -!- FUNCTION: Plays an important role in growth control. Its major
   role in stimulating body growth is to stimulate the liver and
   other tissues to secrete IGF-1. It stimulates both the
   differentiation and proliferation of myoblasts. It also stimulates
   amino acid uptake and protein synthesis in muscle and other
   tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation
   at the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announcement/
   or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U02929; AAA21027.1; .
DR HSSP: P01246; IBS7.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189
FT DISULFID 206 214
SQ SEQUENCE 216 AA; 24423 MW; 37AB3173834D11AC CRC64;
Query Match 64.4%; Score 302.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 2e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;
QY 2 FPTPLSRFDNAMLRAHRLAQAFDYQFEERAYIPKEQKYSFLONPQTSLSFSSESPT 61
DB 27 FPAPPLSLSFANVLRADQLHQLAADTYKEFERAYIPKEQKYS-IQNAQAFCSETIPA 85
QY 62 PSNREETQOKSNLELRISLLLIQSWLEPVQ 92
DB 86 PTGKEAQQRDMELLRFSLIIQSWLGPVQ 116
RESULT 14
SOMA_RABIT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:30 : Search time 35.9427 seconds
(without alignments)
660,520 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: : MFETPLSLRFLDNLAMLRAR.....NLELLRISLLLIQSWLEPVQ 62

Scoring table: BLOSUM62
Gap: 10.0, Gapext 0.5

Searched: 830525 seqs, 238052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database :

SP-REM50.21:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_protist:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirs:
16: sp_bacteriap:
17: sp_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	435	92.6	217	5	Q8WNE0
2	422	89.8	245	4	G14544
3	397	84.5	217	5	Q07369
4	396	84.3	212	6	Q07368
5	396	84.3	217	6	Q07367
6	381	81.1	217	4	G14507
7	348	74.0	217	6	Q8WNE9
8	336.5	71.6	202	4	G14543
9	318	67.7	217	6	Q8M174
10	306.5	65.2	216	11	Q76615
11	301.5	64.1	216	6	Q8M173
12	301.5	64.1	216	6	Q8H555
13	298.5	63.5	216	11	Q9R2C3
14	297.5	63.3	204	5	Q55205
15	297.5	63.3	216	11	Q9JKM4
16	297	63.2	217	6	Q8M175

ALIGNMENTS

RESULT 1

Q8WNE0
ID Q8WNE0 PRELIMINARY: PRI: 217 AA.
AC Q8WNE0
CT Q1-VAR-2002 (TREMHLrel, 20, Created)
DI Q1-VAR-2002 (TREMHLrel, 20, Last sequence update)
DE Q1-VAR-2003 (TREMHLrel, 23, Last annotation update)
DE Growth hormone.
GN GH-N.
OS Ateles geoffroyi (Black-handed spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=9599;
RN 111
RP SEQUENCE FROM N.A.
RA Revd. A., Esquivel D., Santiago D., Barreira-Saldana H.;
RT Independent duplication of the growth hormone gene in three
RT Anthropoid lineages.
RI Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RI EMBL: AF374234; AAL72286.1;
LR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SC SEQUENCE 217 AA; 24894 MW; 425829PF41EEAAE6 CRC64;

Query Match 92.6%; Score 435; DB 6; Length 217;
Best Local Similarity 92.3%; Pred. No. 3.5e-41;
Matches 84; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

CY 2 FFTPLSLRFLDNLAMLRARHLAFLDTVOEPEEYIPKQKYSFQNPOTSLSTFSSEPT 61
|||||
E0 27 FFTPLSLRDLNMLRAHLAFLDTVOEPEEYIPKQKYSFQNPOTSLSTFSSEPT 86
CY 62 PNNRRTTCKSNLELLRISLLLIQSWLEPVQ 92
|||||
D0 87 PASKKETOOKSNLELLRISLLLIQSWLEPVQ 117

Q9tu21 capra hircu
Q9tqw9 bos indicus
Q9fkg0 cavia porce
Q9smj5 tarsius ba
Q9bec9 tragulus ja
P78451 homo sapien
Q9smj6 tarsius syr
Q28957 sus scrofa
Q95240 canis fami
Q9pu72 cynops pyrr
Q14406 homo sapien
Q91386 amia calva
Q8bxx8 bos mutus g
Q8axx9 anguilla an
Q8qfmb ciarias bat
Q98sr8 carassius a
Q98sr8 megalobrama
Q98lt4 megalobrama
Q8qgl1 heteropneus
Q8qfn2 ciarias gar
Q90201 mylopharyng
Q10556 hypophthalm
Q8uvc8 megalobrama
Q8avb2 silurus aso
Q30w26 carassius a
Q8axx5 carassius a
Q98sr7 cyprinus ca

17 290.5 61.8 192 6 Q9TU21
18 289.5 61.6 192 6 Q9TQW9
19 287.5 61.2 190 11 Q9JKG0
20 286.5 61.0 178 6 Q9SMJ5
21 285.5 60.7 217 6 Q9BEC0
22 285.5 60.7 217 6 Q9BEB9
23 285 60.6 167 4 P78451
24 283.5 60.3 178 6 Q9SMJ6
25 280.5 59.7 217 6 Q28957
26 265.5 56.5 143 6 Q95240
27 255.5 54.4 218 13 Q9PU72
28 242.5 51.6 145 6 Q9BDR4
29 234 49.8 199 4 Q14406
30 233.5 49.7 195 13 Q91386
31 186.5 39.7 93 6 Q8BXV8
32 177.5 37.8 209 13 Q8AXX9
33 167.5 35.6 200 13 Q8QFM8
34 166.5 35.4 188 13 Q90283
35 165.5 35.2 187 13 Q98SR8
36 165.5 35.2 188 13 Q98TT4
37 165.5 35.2 200 13 Q8GG11
38 165.5 35.2 200 13 Q8QFN2
39 165.5 35.2 210 13 Q90201
40 165.5 35.2 210 13 Q91056
41 165.5 35.2 210 13 Q8UVG8
42 164.5 35.0 200 13 Q8AV82
43 163.5 34.8 188 13 Q9CM26
44 162.5 34.6 210 13 Q8AXS5
45 161.5 34.4 188 13 Q98SR7

RESULT 2

Q014544
 ID Q014544 PRELIMINARY: PRT: 245 AA.
 AC Q014544
 DT 01-JAN-1998 (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE Placental growth hormone isoform bgh-v3 precursor.
 GN HGH-V
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=98373737; PubMed=9709363;
 RA Boguszewski C.L., Svensson P.A., Jansson T., Clark R.,
 RA Carlsson L.M.S., Carlsson A.
 RT "Cloning of two novel growth hormone transcripts expressed in human
 placenta."
 RL J. Clin. Endocrinol. Metab. 83:2878-2885(1996).
 DR EMBL: AF006061; AAB71829.1;
 DR HSP: P01241; IAZ2.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR Signal.
 FT SIGNAL.
 KW SIGNAL.
 SQ SEQUENCE 245 AA: 27101 MW: 140078CD75D9158 CRC64:
 1 26 POTENTIAL.
 245
 Query Match 89.8%; Score 422; DB 4; Length 245;
 Best Local Similarity 92.3%; Pred. No. 1.2e-39;
 Matches 84; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EPTPLSLFONAMLRARHLHQAFDTYQEFEEAYIPKEQKYSFLQNPOTSLSESIPTP 61
 DB 27 EPTPLSLFONAMLRARHLHQAFDTYQEFEEAYIPKEQKYSFLQNPOTSLSESIPTP 86
 QY 62 PSNREETQOKSNLELLRISILLIQSWLEPVQ 92
 DB 87 PSNRYVTQOKSNLELLRISILLIQSWLEPVQ 117

RESULT 3

Q007369
 ID Q007369 PRELIMINARY: PRT: 217 AA.
 AC Q007369
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Chorionic somatomammotropin-3.
 CS Macaca mulatta (Rhesus macaque).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OX NCBI_TaxID=9544;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Midpregnancy placenta;
 RX MEDLINE=94008724; PubMed=8404617;
 RA Golos T.G., Durning M., Fisher J.M., Fowler P.D.,
 RT "Cloning of four growth hormone/chorionic somatomammotropin-related
 complementary deoxyribonucleic acids differentially expressed during
 pregnancy in the rhesus monkey placenta."
 RL Endocrinology 133:1744-1752(1993).
 DR EMBL: U16554; AAA18841.1;
 DR HSP: P01241; IAXI.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 DR Signal.
 KW NON-TER.
 SQ SEQUENCE 217 AA: 24874 MW: 1106255656F5 CRC64:
 24874
 Query Match 54.3%; Score 396; DB 6; Length 212;
 Best Local Similarity 82.2%; Pred. No. 8.7e-37;
 Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 EPTPLSLFONAMLRARHLHQAFDTYQEFEEAYIPKEQKYSFLQNPOTSLSESIPTP 62
 DB 23 PSVPLSLFONAMLRARHLHQAFDTYQEFEEAYIPKEQKYSFLQNPOTSLSESIPTP 82
 QY 63 SNREETQOKSNLELLRISILLIQSWLEPVQ 92
 DB 83 SNLEETQOKSNLELLRISILLIQSWLEPVQ 112
 RESULT 5
 Q007367
 ID Q007367 PRELIMINARY: PRT: 217 AA.
 AC Q007367
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Chorionic somatomammotropin-1.
 CS Macaca mulatta (Rhesus macaque).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OX NCBI_TaxID=9544;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Midpregnancy placenta;

[illegible]

```

RA Boquezewski C.L., Svensson P.A., Jansson E., Clark R.
RA Carlsson L.M.S., Carlsson B.
RT "Cloning of two novel growth hormone transcripts expressed in human
RT placenta."
RL J. Clin. Endocrinol. Metab. 83:2878-2886(1998).
DR FMBL; AF060600; AAR1828.1; -.
DR HSSP; P01241; I422.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PR00836; Somatotropin.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL
SQ SEQUENCE 202 AA; 23128 MW; 36B64D05A9197C6 CRC64;

Query Match 71.6%; Score 336.5; DP 4; Length 202;
Best Local Similarity 76.9%; Pred. No. 4.4e-30;
Matches 70; Conservative 3; Mismatches 3; Indels 15; Gaps 2;

QY 2 PPTPLSLFDNANLRAHRLHQLAFDYQEFEEAYIPKQYSLQNPQISLFSSESIPT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 PPTPLSLFDNANLRAHRLHQLAFDYQEF-----NPOISLFSSESIPT 71

QY 62 PSNRETQKSNLELRISILLIQSWLEPVQ 92
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72 PSNRETQKSNLELRISILLIQSWLEPVQ 102

RESULT 9
Q8MI74
ID Q8MI74 PRELIMINARY: PRT: 217 AA.
AC Q8MI74;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Growth hormone-like protein 6 precursor.
GN GHLP6.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrich.
OX NCBI_TaxID=9483;
RN [1]
RA SEQUENCE FROM N.A.
RP Wallis O.C., Wallis M.
RT "Characterisation of the GH gene cluster in a New world monkey, the
RT marmoset (Callithrix jacchus).";
RE J. Mol. Endocrinol. 5:0-0(2002).
DR FMBL; AJ489811; CA334012.1; -.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PR00836; Somatotropin.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL
FT CHAIN 1 26 POTENTIAL.
SQ SEQUENCE 217 AA; 25177 MW; 5ECF1A875E278F1A CRC64;

Query Match 67.7%; Score 318; DP 4; Length 217;
Best Local Similarity 68.9%; Pred. No. 5.9e-28;
Matches 62; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 3 PPTPLSLFDNANLRAHRLHQLAFDYQEFEEAYIPKQYSLQNPQISLFSSESIPT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
28 PPTPLSLFDNANLRAHRLHQLAFDYQEFKNCVPRKQYFLNRPETVFSSESIPT 67

QY 63 PSNRETQKSNLELRISILLIQSWLEPVQ 92
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 PSNRETQKSNLELRISILLIQSWLEPVQ 117

RESULT 10
Q70615
ID Q70615 PRELIMINARY: PRT: 216 AA.
AC Q70615;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Growth hormone precursor.
OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
OC Namnospalax.
OX NCBI_TaxID=30637;
RN [1]
RA SEQUENCE FROM N.A.
RP MED; NE-99124645; PubMed-9924177;
RA Lioupis A., Nevo E., Wallis M.
RT "Cloning and characterisation of the gene encoding mole rat (Spalax
RT ehrenbergi) growth hormone";
RL J. Mol. Endocrinol. 22:29-36(1999).
DR FMBL; AJ005819; CA06716.1; -.
DR HSSP; P01241; I4X1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PR00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL
FT CHAIN 1 26 POTENTIAL.
FT SIGNAL 27 216 GROWTH HORMONE.
SQ SEQUENCE 216 AA; 24627 MW; EEAB8A523BA0A0FE CRC64;

Query Match 65.2%; Score 306.5; DP 11; Length 216;
Best Local Similarity 65.9%; Pred. No. 1.2e-26;
Matches 60; Conservative 13; Mismatches 17; Indels 1; Gaps 1;

QY 2 PPTPLSLFDNANLRAHRLHQLAFDYQEFEEAYIPKQYSLQNPQISLFSSESIPT 61
Ub ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 PPTPLSLFDNANLRAHRLHQLAFDYQEFKNCVPRKQYFLNRPETVFSSESIPT 85

QY 62 PSNRETQKSNLELRISILLIQSWLEPVQ 92
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86 PTKSEAKQSRSDMELIRFSLILIQSWLGPVQ 116

RESULT 11
Q8MI73
ID Q8MI73 PRELIMINARY: PRT: 216 AA.
AC Q8MI73;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Growth hormone precursor.
GN GH.
OS Delphinus delphis (Saddleback dolphin) (Black sea dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Delphinus.
OX NCBI_TaxID=9728;
RN [1]
RA SEQUENCE FROM N.A.
RP L'SSEE=Liver;
RA Kanion Z., Wallis O.C., Wallis M.
RT "Cloning and characterisation of the GH gene from the common dolphin
RT (Delphinus delphis).";
RE submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR FMBL; AJ49191; CAD37292.2; -.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PR00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL
FT CHAIN 1 26 POTENTIAL.
FT SIGNAL 27 216 GROWTH HORMONE.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:44:15 : Search time 60.4158 seconds
(without alignments)
225.942 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 FVNHLCGSHVEP-LVCG.....IVFQCTICSLYQLENCHN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107663 seqs, 158724573 residues

Total number of hits satisfying chosen parameters: 1107663

Minimum DB seq length: 0
Maximum DB seq length: 2003000090

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_13Jun03.*
1: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1960.DAT.*
2: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1961.DAT.*
3: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1962.DAT.*
4: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1963.DAT.*
5: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1964.DAT.*
6: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1965.DAT.*
7: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1966.DAT.*
8: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1967.DAT.*
9: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1968.DAT.*
10: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1969.DAT.*
11: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1970.DAT.*
12: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1971.DAT.*
13: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1972.DAT.*
14: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1973.DAT.*
15: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1974.DAT.*
16: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1975.DAT.*
17: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1976.DAT.*
18: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1977.DAT.*
19: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1978.DAT.*
20: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1979.DAT.*
21: /SIDSL/qcdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/qcdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/qcdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/qcdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	86	5	AAP40829
2	463	100.0	86	17	AAR84061
3	463	100.0	86	20	AA142858
4	463	100.0	86	21	AA12770
5	463	100.0	96	23	AA143218
6	463	100.0	97	3	AAP20036
7	463	100.0	87	5	AAP40217
8	463	100.0	87	5	AAP50127
9	463	100.0	87	5	AAP50063

10	463	100.0	87	7	AAP61090
11	463	100.0	87	14	AAR32357
12	463	100.0	88	11	AAR07682
13	463	100.0	88	14	AAR33855
14	463	100.0	92	13	AAR20467
15	463	100.0	93	10	AAP50102
16	463	100.0	96	20	AA017830
17	463	100.0	96	20	AA108004
18	463	100.0	97	15	AAR68898
19	463	100.0	97	21	AA12773
20	463	100.0	110	2	AAP10042
21	463	100.0	110	2	AAP10053
22	463	100.0	110	5	AAP40309
23	463	100.0	110	20	AA106608
24	463	100.0	110	21	AAB26765
25	463	100.0	110	21	AAB06144
26	463	100.0	110	21	AA170366
27	463	100.0	110	21	AA144367
28	463	100.0	110	22	AAG65677
29	463	100.0	110	22	AAE10337
30	463	100.0	110	22	AAB35424
31	463	100.0	110	23	ABG31590
32	463	100.0	110	23	ABG50634
33	463	100.0	117	21	AA159788
34	463	100.0	130	21	AA169787
35	463	100.0	151	20	AAW81856
36	463	100.0	153	21	AA133589
37	463	100.0	161	23	AAW81857
38	457	99.7	88	14	AAR35574
39	457	99.7	88	14	AAR35575
40	457	99.7	88	14	AAR35576
41	457	99.7	96	20	AA017831
42	457	99.7	96	20	AA108005
43	457	99.7	97	15	AAR78661
44	456	98.5	98	21	AA151222
45	456	98.5	99	21	AA151223

ALIGNMENTS

FASTA :
AAP40829
10 AAP40829 standard; Protein: 86 AA.
XX
XX AAP40829;
NC
CT 25-MAR-2003 (updated)
CT 03-AUG-1992 (first entry)
XX
DE
XX Sequence of human insulin precursor.
XX Insulin precursor; connecting peptide; diabetes; hormone.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Region 1..30
XX Peptide 31..65
XX Region 56..86
XX Modified-site 1
XX /label= chain B
XX /label= connecting peptide
XX /label= chain A
XX /label= F-NH2-R
XX /note= "H or a chemically or enzymatically cleavable AA residue or peptide residue"
XX Disulfide bond 7..72
XX Disulfide bond 19..85
XX Disulfide bond 71..76
XX Modified-site 86
XX /label= N-OH

Sequence encoded b
Proinsulin protein
Modified human ins
hpl. Homo sapiens
Yeast alpha-factor
Synthetic proinsul
Human proinsulin #
Human proinsulin p
Human pro-insulin
Human proinsulin p
Sequence encoded b
Sequence of prepro
Sequence of human
Human preproinsuli
Human preproinsuli
Human insulin. Ho
Human insulin used
Human proinsulin.
Human proinsulin n
Secretory cell lin
Human preproinsuli
Human insulin. Ho
MKPsp-KMPsp10-Met-
MKPsp-KMPsp10-(His
Human proinsulin a
Human preproinsuli
Human proinsulin a
Human Met-Tyr pro
Human Met-Phe pro
Human proinsulin #
Human proinsulin p
Proinsulin sequenc
Human preproinsuli
Human preproinsuli

PN US4430266-A.
 XX
 PC 07-FEB-1984.
 XX
 PF 16-FEB-1982; 82US-0349397.
 XX
 PR 16-FEB-1982; 82US-0349397.
 PR 27-MAR-1982; 82US-0134869.
 PR 28-NOV-1982; 82US-0210696.
 XX
 PA (SULL) DILLY & CO. ELL.
 XX
 PI Frack 2H;
 XX
 PI WPI: 1984-349032/08.
 XX
 XX Insulin precursor prodm. from linear S-sulphonate and mercaptan -
 PT in single step without separate oxidn.
 XX
 PS Claim 17: Column 4; 8pp; English.
 XX
 CC The inventors claim a method for the prepn. of an insulin precursor
 CC in which the A-chain and B-chain are joined through a connecting
 CC peptide. The connecting peptide joins the A-chain at the amino
 CC group of A-1 to the B-chain at the carboxyl group of B-30. The
 CC method is pref. for the prepn. of human insulin precursor (see
 CC AAP40825). The SOS of the connecting peptides of a number of species
 CC are given (see AAP40824, AAP40830-33).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 86 AA;
 Query Match 100.0%; Score 463; DB 5; Length 86;
 Best Local Similarity 100.0%; Pred. No. 8.4e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGPGAGSLQPLALEG 60
 DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGPGAGSLQPLALEG 60
 QY 61 SLCKRGIVECCISCSLYOLENYCN 86
 DB 61 SLCKRGIVECCISCSLYOLENYCN 86
 RESULT 2
 AAR84061
 ID AAR84061 standard; Protein: 86 AA.
 AC AAR84061;
 XX 22-AUG-1996 (first entry);
 XX Human insulin.
 XX Insulin; transformation; gene expression; fungal; fungal cells;
 KW hormone; A-chain; C-chain; glycosylation.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 1..26; a
 FT /tag= a
 FT /product. Insulin.
 XX
 XX EP704527-A.
 XX
 XX 03-APR-1996.
 XX 03-AUG-1995; 95EP-0001399.
 XX 05-AUG-1994; 94HR-0003432.
 XX

PA (WEST/) MESTRIC S.
 XX (PLIV) PLIVA PHARM & CHEM FAB.
 XX
 XX Mestric S, Punt PJ, Valinger R, Van DEN HONDEL CAM33;
 XX
 XX WPI: 1996-129917/14.
 XX
 XX N-PS28; AAP17830, AAP17831.
 XX
 XX DNA encoding human insulin precursors - which comprise B- and
 XX A-chains linked via amino acid chain confg; 1 of core glycosylation
 XX sites; for prepn. of insulin in fungal cells
 XX
 XX Disclosure: Figure 1; 32pp; English.
 XX
 XX LNA sequences encoding insulin precursors of formula H-Pq-A, where B
 XX and A represent B- and A-chains of insulin respectively, and Pq
 XX represents a modified C-peptide or any number of amino acids
 XX comprising at least one glycosylation consensus site, can be
 XX inserted into expression vectors which in turn can be used to
 XX transform fungal host cells. The fungal cells are then cultured
 XX and the insulin expressed in such cells can be harvested.
 XX
 SQ Sequence 86 AA;
 Query Match 100.0%; Score 463; DB 17; Length 86;
 Best Local Similarity 100.0%; Pred. No. 8.4e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGPGAGSLQPLALEG 60
 DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGPGAGSLQPLALEG 60
 QY 61 SLCKRGIVECCISCSLYOLENYCN 86
 DB 61 SLCKRGIVECCISCSLYOLENYCN 86
 RESULT 3
 AAY42558
 ID AAY42558 standard; Protein: 86 AA.
 AC AAY42558;
 XX 19-JAN-2000 (first entry)
 XX Human insulin precursor, SEQ ID 5.
 XX Insulin; precursor; growth hormone; chaperone; intramolecular;
 XX folding; conformation; chimeric protein; cleavable; recombinant;
 XX production; yield.
 XX Homo sapiens.
 XX W05950302-A1.
 XX 07-OCT-1999.
 XX 31-MAR-1998; 98WO-CN00052.
 XX 31-MAR-1998; 98WO-CN00052.
 XX (TONG) LONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX Gan Z;
 XX WPI: 1999-610834/52.
 XX New chimeric proteins containing human growth hormone fragment, used
 XX particularly for the production of human insulin -
 XX Claim 10: Page 29; 46pp; English.
 XX This sequence represents a human insulin precursor comprising insulin

CC A and B chains separated by a 34 residue peptide sequence. This insulin
 CC precursor can be a component of chimeric proteins which additionally
 CC contains an N-terminal fragment of human growth hormone (hGH); and a
 CC cleavable peptide linker (RAY42857). The hGH portion of the chimeric
 CC protein acts as an intramolecular chaperone (IMC) for the insulin
 CC precursor, enabling it to fold correctly. The cleavable peptide linker
 CC has a C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimera protein can
 CC provide human insulin with correctly linked cysteine bridges with
 CC fewer necessary procedural steps, and hence resulting in a higher yield
 CC of human insulin. The IMC sequences not only protect insulin sequences
 CC from intracellular degradation by a microorganism host, but also promote
 CC the folding of the fused insulin precursor, facilitate the solubility of
 CC the fusion protein and decrease the intermolecular interactions among
 CC the fusion proteins, thus allowing folding of the fused insulin precursor
 CC at commercially useful high concentrations. The procedural steps of
 CC cyanogen bromide cleavage, oxidative sulphatolysis and related
 CC purification steps can thus be eliminated, along with the use of high
 CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
 XX Sequence 86 AA:

Query Match 100.0%; Score 463; DP 20; Length 86;

Best Local Similarity 100.0%; Pred. No. 8,4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVYALYVCGRGFFYTPKTRPEADQLQGVQLGSGPGAGS;QPLALRG 60

DB 1 FVNHLCGSHLVYALYVCGRGFFYTPKTRPEADQLQGVQLGSGPGAGS;QPLALRG 60

QY 61 SLQKRGIVECCTSTCSLYOLENYCN 86

DB 61 SLQKRGIVECCTSTCSLYOLENYCN 86

QY 61 SLQKRGIVECCTSTCSLYOLENYCN 86

DB 61 SLQKRGIVECCTSTCSLYOLENYCN 86

RESULT 4

ID AAB12770

XX AAB12770 standard; protein: 86 AA.

AC AAB12770;

XX 22-NOV-2000 (first entry)

XX Human proinsulin protein sequence SEQ ID NO:2.

XX Human; insulin-like growth factor 1; IGF-1; Proinsulin; insulin.

KW mutant; variant; insulin-like growth factor binding protein; IGFBP-1;

KW IGFBP-3; antidiabetic; neuroprotective; anorectic; tranquiliser;

KW vulnery; anorectic; cardiatic; nephrotropic; dermatological; antiHIV;

KW antiviral; hyperglycaemia; obesity; lung disease; glomerulonephritis;

KW interstitial nephritis; Turner's syndrome; Laron's syndrome;

KW short stature; increased fat mass-to-lean ratio; immunological disorder;

KW peripheral neuropathy; multiple sclerosis; muscular dystrophy;

KW catabolic state; trauma; wounding; infection; HIV; skin disorder;

KW human immunodeficiency virus; diabetes; heart dysfunction;

KW kidney disorder; whole body growth disorder.

XX Homo sapiens.

OS WC2000040612-A1.

XX 13-JUL-2000.

XX 05-JAN-2000; 2000WO-US00151.

XX 06-JAN-1999; 99US-0115010.

XX (GETH) GENENTECH INC.

XX Dubaquié Y, Lowman HB;

XX WPI: 2000-455955/40.

XX Novel insulin-like growth factor (IGF) 1 mutants that selectively bind
 PT to IGF binding protein (IGFBP)-1 or IGFBP-3, used to improve the
 PT half-lives of IGF-1 and insulin -

XX Disclosure: Page 44; 48pp; English.

XX The present invention describes an insulin-like growth factor (IGF)-1
 CC variant (I), where an amino acid at position 3, 4, 5, 7, 10, 14, 17, 23,
 CC 24, 25, 43, 49 or 53, optionally in combination with an amino acid at
 CC position 12 and/or 16 of the native human IGF-1 sequence, is replaced
 CC with an alanine, glycine, or a serine residue. The residue at position 7
 CC may be replaced by any amino acid. (I) can have antidiabetic, cardiac,
 CC neuroprotective, anorectic, tranquiliser, vulnery, anorectic,
 CC nephrotropic, dermatological, antiHIV and antiviral activities. The IGF-1
 CC mutants are used in any methods where IGFs or insulin are used, e.g. in
 CC treating hyperglycaemia, obesity-related, neurological, cardiac, renal,
 CC immunological, and anabolic disorders. These disorders include lung
 CC diseases, glomerulonephritis, interstitial nephritis, Turner's syndrome,
 CC Laron's syndrome, short stature, increased fat mass-to-lean ratios,
 CC immunological disorders, peripheral neuropathy, multiple sclerosis,
 CC muscular dystrophy, catabolic states, trauma, wounding, infection,
 CC human immunodeficiency virus (HIV), wounds, skin disorders, diabetes,
 CC heart dysfunctions, kidney disorders, and whole body growth disorders.
 CC They can also be used for increasing serum and tissue levels of
 CC biological active IGF or insulin in a mammal. The IGF-1 mutants improve the
 CC half-lives of IGF-1 and insulin. The present sequence represents the
 CC native human proinsulin protein sequence, which is given in the
 CC exemplification of the present invention.

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DP 20; Length 86;

Best Local Similarity 100.0%; Pred. No. 8,4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVYALYVCGRGFFYTPKTRPEADQLQGVQLGSGPGAGS;QPLALRG 60

DB 1 FVNHLCGSHLVYALYVCGRGFFYTPKTRPEADQLQGVQLGSGPGAGS;QPLALRG 60

QY 61 SLQKRGIVECCTSTCSLYOLENYCN 86

DB 61 SLQKRGIVECCTSTCSLYOLENYCN 86

QY 61 SLQKRGIVECCTSTCSLYOLENYCN 86

DB 61 SLQKRGIVECCTSTCSLYOLENYCN 86

RESULT 5

ID AAM48218

XX AAM48218 standard; protein: 86 AA.

AC AAM48218;

XX 16-MAR-2002 (first entry)

XX Human proinsulin.

XX Actirheumatic; antiarthritic; osteopathic; cartilage disorder;

KW insulin-like growth factor; IGF; binding protein; IGFBP;

KW rheumatoid arthritis; osteoarthritis; proinsulin; human.

XX Homo sapiens.

OS WC200187323-A2.

XX 22-NOV-2001.

XX 16-MAY-2001; 2001WO-US15934.

XX 16-MAY-2000; 2000US-204490F.

XX 15-NOV-2000; 2000US-248985P.

XX (GETH) GENENTECH INC.

XX Dubaquié Y, Filvarolf EH, Lowman HB;

```

XX DR WPI: 2002-082942/11.
XX PT Treating cartilage disorders including cartilage damage by injury or
XX PT degenerative cartilaginous disorders, by contacting cartilage with
XX PT insulin-like growth factor analog with altered affinity for IGF-binding
XX PT proteins.
XX PS Disclosure: Fig 16: 136pp; English.
XX CC The present invention relates to a method for treating cartilage
XX CC disorders. The method comprises contacting cartilage with an active agent
XX CC such as insulin like growth factor (IGF-1) analog with a binding affinity
XX CC preference for IGF binding protein-3 (IGFBP-3) over IGFBP-1 and an IGF-1
XX CC analog with a binding affinity preference for IGFBP-1 over IGFBP-3, or a
XX CC IGFBP displacer peptide that prevents the interaction of IGF with an
XX CC IGFBP and does not bind to human IGF receptor. The method is useful for
XX CC treating cartilage disorders (CD), including degenerative CD, articular
XX CC CD such as rheumatoid arthritis and osteoarthritis. The present sequence
XX CC is human proinsulin, which was used to illustrate the invention.
XX SQ Sequence 85 AA:
    Query Match: 100.0%; Score 463; DB 23; Length 86;
    Best Local Similarity: 100.0%; Pred. NO. 8.4e-43;
    Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 FVNQLCGSHLVEA:YLVCGERGFFYPKTRREAEDLVQGVGGPGAGSGPLALEG 60
    DB 1 FVNQLCGSHLVEA:YLVCGERGFFYPKTRREAEDLVQGVGGPGAGSGPLALEG 60
    QY 61 SLQKRGIVEQCCTSIQSIYQLENYCN 86
    DB 61 SLQKRGIVEQCCTSIQSIYQLENYCN 86
    RESULT 6
    AAP20036 standard: Protein: 87 AA.
    AC AAP20036;
    DT 25-MAR-2003 (updated)
    DI 22-JUL-1992 (first entry)
    DE Human proinsulin.
    DE Proinsulin.
    DE Homo sapiens.
    PN EP55942-A.
    PD 14-JUL-1982.
    PF 31-DEC-1982; 81EP-0306129.
    PR 02-JAN-1981; 81US-0222010.
    PR 23-JUL-1981; 81US-0286076.
    PR 02-JAN-1982; 82US-0222010.
    PR 03-MAR-1982; 82US 0354287.
    XX (CNYN-) STATE UNIV NEW YORK.
    XX Inouye M, Nakamura K;
    DR WPI: 1982-59775E/29 (59775E).
    DR N-PSDB; AAN20041.
    XX Plasmid cloning vehicles - useful for transforming bacterial
    PT hosts to produce eukaryotic polypeptide(s).
    XX Disclosure: Fig 27: 114pp; English.

```

```

XX CC The sequence comprises human proinsulin.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX SQ Sequence 87 AA:
    Query Match: 100.0%; Score 463; DB 3; Length 87;
    Best Local Similarity: 100.0%; Pred. NO. 8.5e-43;
    Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 FVNQLCGSHLVEA:YLVCGERGFFYPKTRREAEDLVQGVGGPGAGSGPLALEG 60
    DB 1 FVNQLCGSHLVEA:YLVCGERGFFYPKTRREAEDLVQGVGGPGAGSGPLALEG 61
    QY 61 SLQKRGIVEQCCTSIQSIYQLENYCN 86
    DB 62 SLQKRGIVEQCCTSIQSIYQLENYCN 87
    RESULT 7
    AAP40217 standard: Protein: 87 AA.
    AC AAP40217;
    DT 25-MAR-2003 (updated)
    DI 12-FEB-1992 (first entry)
    DE Sequence of the 32 N-terminal AAs of proinsulin.
    DE Hormone; cloning vector; phage resistant.
    OS Homo sapiens.
    CH Key
    FT Region
    FT /label= B-chain
    FT Region
    FT /label= C-chain
    FT Region
    FT /label= A-chain
    XX GB2126237-A.
    XX 21-MAR-1984.
    XX 01-SEP-1983; 83GS-0023468.
    XX 03-SEP-1982; 82US-0414290.
    XX 05-SEP-1984; 84US-0647338.
    XX (EELI) LILLY & CO ELI.
    XX Hershberger CL, Rostock PR;
    DE WPI: 1984-070793/12.
    DE N-PSDB; AAN40179.
    XX Protecting bacteria from phage infection - by transformation with
    FT cloning vector contg. segment with restriction and modification
    FT activity
    XX Example: Fig 10: 28pp; English.
    XX Plasmid pTh alpha 1 was constructed by inserting a synthesised gene
    CC for thymosin alpha 1 (AAN40178) into plasmid pBR322. It is used for
    CC the construction of pIrp24. The inventors claim a method for
    CC protecting bacteria from phage infection - by transformation with
    CC cloning vector contg. segment with restriction and modification
    CC activity. Prodn. of plasmid pPR 26 or pPR27 which uses pIrp24; and
    CC prodn. of plasmid pPR29 which uses a synthetic gene coding for the
    CC 32 N-terminal AAs of proinsulin (see AAN40179).
    CC (Updated on 25-MAR-2003 to correct PA field.)

```

```

XX SQ Sequence 87 AA:
    Query Match 100.0%; Score 463; DB #: Length 87;
    Best Local Similarity 100.0%; Pred. No. 8.5e-43;
    Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAERLQVGVEVGSGGASLSLOFLAER 60
    DQ 1 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAERLQVGVEVGSGGASLSLOFLAER 60
    QY 2 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAERLQVGVEVGSGGASLSLOFLAER 60
    DQ 2 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAERLQVGVEVGSGGASLSLOFLAER 60

QY 51 SLOKRGIVEQCCTSIQSLYOLENYCN 86
    DQ 51 SLOKRGIVEQCCTSIQSLYOLENYCN 86
    QY 62 SLOKRGIVEQCCTSIQSLYOLENYCN 87
    DQ 62 SLOKRGIVEQCCTSIQSLYOLENYCN 87

RESULT 8
AAP50127
ID AAP50127 standard: Protein: 67 AA.
XX
AC AAP50127;
XX
DT 25-MAR-2003 (updated)
DT 16-AUG-2002 (updated)
DT 30-SEP-1991 (first entry)
XX
DE Sequence of the 32 N-terminal AAs of proteinase.
XX
KW Selectable vector; autonomously replicating vector;
KW expression vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region: 2..31
FT Region: /label= A chain
FT Region: 32..66
FT Region: /label= B chain
FT Region: 67..87
FT Region: /label= A chain
XX
PN EPI54539-A.
XX
PD 11-SEP-1985.
XX
PF 04-MAR-1985; 8SEP-0301452.
XX
PR C6-MAR-1984; 84US-0586592.
PR 01-APR-1985; 85CN-010555.
XX
PA (EPI ) LILLY & CO ELL.
XX
PI Schoner R, Schoner R;
XX
WP1; 1985-224921/37.
DR N-PSDS; AAN50152.
XX
PT New recombinant DNA expression vector with autonomous
FT replication and on transcription generating polycistronic mRNA
XX
PS Example; Fig 14; 118pp; English.
XX
CC The inventors claim a process for preparing selectable and
CC autonomously replicating recombinant DNA expression vectors which
CC comprise 1) a transcriptional and translational activating sequence
CC which is in the reading frame of a nucleotide sequence which codes
CC for a peptide or polypeptide; 2) a translational stop signal; 3) a
CC translational start signal which is in the reading frame of a
CC nucleotide sequence that codes for a functional polypeptide;
CC and 4) an additional translational stop signal. The peptide or
CC polypeptide coding sequence codes for 2-20 AAs, esp. AAP50122-P50125.
CC The functional polypeptide is esp. growth hormone, human insulin,

```

```

CC Interferon and human tissue plasminogen activator.
CC (Updated on 16-AUG-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SC Sequence 87 AA;

Query Match 100.0%; Score 463; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 8.5e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHCGSLHVEALYLVCGGCGFFYTPKTRREAEIVQGVGVGGGAGSLOPLAEG 60
DB 2 FVNOHCGSLHVEALYLVCGGCGFFYTPKTRREAEIVQGVGVGGGAGSLOPLAEG 61
QY 61 SLOKRGIVECCCTSCSLYQLENYCN 86
DB 62 SLOKRGIVECCCTSCSLYQLENYCN 87
QY 61 SLOKRGIVECCCTSCSLYQLENYCN 86
DB 62 SLOKRGIVECCCTSCSLYQLENYCN 87

RESULT 9
AAP50060
LC AAP50060 standard; protein; 87 AA.
XX
AC AAP50060;
XX
XX 25-MAR-2003 (updated)
ET 16-AUG-2002 (updated)
ET 11-NOV-1991 (first entry)
XX
DE Synthetic proinsulin.
XX
XX Proinsulin; vector; proteinaceous granule;
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..30 /label=B chain.
XX Region 31..65 /label=C chain.
XX Region 66..86 /label=A chain.
XX
XX EP159123.A.
XX
XX 23-OCT-1985.
XX
XX 04-MAR-1985; 85EP-0301459.
XX
XX 06-MAR-1984; 84JUS-0586592.
XX 26-JUL-1984; 84JUS-0634920.
XX 31-JAN-1985; 85JUS-0697090.
XX
XX (EJL) Lilly & Co. Eli.
XX
XX Hsiung HM, Schiner RG, Schoner BE;
XX
XX WPI: 1985-265030/43.
XX N-PSDB: AAN50082.
XX
XX New selectable and autonomously replicating DNA expression vector -
XX Useful in producing proteinaceous granules in cell transformants, esp.
XX
XX Disclosure; Fig. 14; 115pp; English.
XX
XX The synthetic proinsulin gene is expressed in a new selectable and
XX autonomously replicating recombinant DNA expression vector
XX comprising a runaway replicon and a transcriptional and
XX translational activating sequence in the reading frame of the
XX proinsulin coding sequence, the sequence contg. a translational
XX stop signal. Host cells contg. the vector, which is esp. p.asmid
XX pc2103, are cultured, and proinsulin is produced as a highly
XX homo-eous species of proteinaceous granule. The granule can be

```

CC readily isolated from cell lysates and is stable on washing with
 CC urea or detergent soaps, at low concentrations. The aromatic contains at
 CC least 50% of proinsulin and all isolation operations are
 CC simplified.
 CC (Updated on 16-AUG-2002 to add missing GS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 61
 DB 2 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVECCCTSGSLYLENYCN 86
 DB 62 SLOKRGIVECCCTSGSLYLENYCN 87

RESULT 10
 AAP61090 standard; Protein: 87 AA.
 XX
 AC AAP61090;
 XX
 DC 28-FEB-1992 (first entry)
 XX
 DE Sequence encoded by the structural gene for human proinsulin.
 XX
 KW Recombinant plasmid; E.coli expression vector; secretion vector.
 XX
 OS Homo sapiens.
 XX
 PN US4624926-A.
 XX
 PD 25-NOV-1986.
 XX
 XX 03-MAR-1982; 82US-0354287.
 XX
 PR 03-MAR-1982; 82US-0354287.
 PR 02-JAN-1981; 81US-0222010.
 PR 23-JUL-1981; 81US-0286070.
 XX
 XX (UNY-) UNIV OF NEW YORK.
 XX
 XX Inouye M, Nakamura K;
 XX
 DR WPI: 1986-331802/50.
 DR N-PSDB; AAN60872.
 XX
 XX New recombinant plasmid(s) - conts. DNA sequences encoding
 XX exogenous polypeptide and outer membrane protein of E.coli;
 XX
 PS Example: Fig 27; 44pp; English.

XX The inventors claim new recombinant plasmids contain a DNA sequence
 CC encoding a polypeptide, which is foreign to E.coli, in reading phase
 CC with a DNA SQ, coding for at least one functional fragment derived
 CC from an outer membrane lipoprotein gene of E.coli. The foreign gene
 CC may be for human insulin. The lipoprotein gene functional fragment
 CC may be the promoter, the 5'-UTR, the 3'-UTR or the transcription
 CC termination signal provided that it includes at least the promoter.
 XX
 SQ Sequence 87 AA;
 Query Match 100.0%; Score 463; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 61

DB 2 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVECCCTSGSLYLENYCN 86
 DB 62 SLOKRGIVECCCTSGSLYLENYCN 87

RESULT 11
 AAR32367 standard; Protein: 87 AA.
 XX
 AC AAR32367;
 XX
 DT 25-MAR-2003 (updated)
 DT 16-JUN-1993 (first entry)
 XX
 XX Proinsulin protein sequence.
 XX
 KW Human; proinsulin; vector; pUC19; ppINS; CAT; pUC-CAT-proinsulin;
 KW insulin analogue; type I; type II; diabetes.
 XX
 OS Synthetic.

XX
 PN WO9303174-A1.
 XX
 PD 18-FEB-1993.
 XX
 PF 31-JUL-1992; 92WO-0506451.
 XX
 PR 08-AUG-1991; 91US-0741938.
 PR 30-JUL-1992; 92US-0918953.
 XX
 PA (P12) PEIZER INC.
 PA (SCIO) SCIOS INC.
 XX
 PI Andy RJ, Larson EK;
 XX
 DR WPI: 1993-076530/09.
 DR N-PSDB; AAO37003.
 XX
 XX New hepato selective and peripheral selective human insulin
 XX analogues - and their corresp. DNA, for treatment of type I and
 XX type II diabetes
 PS Disclosure: Fig 2b; 58pp; English.
 XX
 XX This sequence represents human proinsulin and was decoded from
 XX the sequences given in AAO36996-7001. The cDNA fragment coding for
 XX proinsulin was inserted into plasmid vector pUC19 and digested with
 XX KpnI and HindIII. This resulted in the formation of the vector pFINS.
 XX A fragment encoding amino acids 1-73 of CAT (see AAO37002) was inserted
 XX into pFINS to give a plasmid which contained DNA sequences which coded
 XX for amino acids 1-73 of CAT, an 8 amino acid linker sequence and human
 XX proinsulin. This plasmid, pUC-CAT-proinsulin, could be used in the
 XX formation of insulin analogues which may be used in the treatment of
 XX types I and II diabetes.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 14; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 60
 DB 2 FVNHLCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVDELGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVECCCTSGSLYLENYCN 86
 DB 62 SLOKRGIVECCCTSGSLYLENYCN 87

RESULT 12

AA07682
ID AAR07682 standard; protein: 86 AA.

XX AC AAR07682;
XX AC
XX 25-MAR-2003 (updated)
DT 09-JAN-2003 (updated)
DT 13-FEB-1991 (first entry)

XX DE Modified human insulin precursor.

XX XX Human insulin precursor; cathepsin C.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers
FT Peptide 1..12

FT Peptide /label-N-terminal initiating dipeptide

FT Peptide 1..32

FT Peptide /label-native human insulin B-chain

FT Peptide 33..67

FT Peptide /label-natural connecting peptide of human

FT Peptide proinsulin

FT Peptide 68..88

FT Peptide /label-native human insulin A-chain

XX EP397420-A.

XX 14-NOV-1990.

XX 04-MAY-1990; 80EP-0304890.

XX 09-MAY-1989; 89US-0349472.

XX (ELIL) LILLY & CO ELI.

XX PI Becker GW, Furman TC, Mackellar WT, McKeough JB;

XX WPI: 1990-343372/45.

XX Human insulin precursor - contg. Met-Tyr or Met-Tyr initiating

PT dipeptide for controlled removal by cathepsin C

XX Disclosure; page 3; 8pp; English.

XX This modified human insulin precursor comprises an N-terminal

CC initiating dipeptide, chosen from Met-Tyr or Met-Tyr, which does

CC not define a cathepsin C dipeptide removal stop point. This

CC dipeptide is linked to the natural human insulin B-chain, natural

CC human proinsulin connecting peptide and natural human insulin A-

CC chain. Dipeptide removal is carefully controlled to obtain the

CC desired prod. without further degradation occurring, irrespective

CC of whether the next dipeptide in the sequence defines a cathepsin

CC C stop point.

CC (Updated on 29-JAN-2003 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.6e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLQGVGLGGPGAGSLQPLALEG 69

DB 3 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLQGVGLGGPGAGSLQPLALEG 62

QY 61 SLOKRGIVECCCTSGISLYOLENYCN 86

DB 63 SLOKRGIVECCCTSGISLYOLENYCN 88

RESULT 13

AA07685
ID AAR07685 standard; Protein: 88 AA.

XX AC AAR07685;

XX AC

XX 25-MAR-2003 (updated)

DT 19-JUL-1993 (first entry)

XX XX hpl.

XX XX

XX Proinsulin; hpl; native; pCZR126S; expression vector; E. coli; human;

XX expression; immunological effect.

XX OS Homo sapiens.

XX XX

XX EP534705-A2.

XX 31-MAR-1993.

XX 22-SEP-1992; 92EP-0308601.

XX 24-SEP-1991; 91US-0764655.

XX (ELIL) LILLY & CO ELI.

XX PI Biogenade RM;

XX XX

XX WPI: 1993-102806/13.

XX N-PSDB; AAQ38310.

XX Expression of low molecular wt. polypeptide(s) e.g. insulin

PT growth factor I - by expressing as deriv. with N-terminal

PT amino acid to provide increased expression levels

XX Disclosure; Page 21-22; 40pp; English.

XX This sequence represents an analogue of native human proinsulin (hpl).

CC The DNA encoding this sequence was used in the construction of the

CC expression vector of the invention. The coding region of the hpl gene

CC was synthesised and was cloned into the expression plasmid pCZR126S

CC (see also AAQ38307). Expression of this gene lead to the inclusion of

CC an extra amino acid (Arg) in the second position from the N-terminal

CC of mature hpl. The extra amino acid provides increased expression

CC levels of the protein and is then cleaved off to avoid undesirable

CC immunological effects when used in humans.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 88 AA;

Query Match 100.0%; Score 463; DB 14; Length 88;

Best Local Similarity 100.0%; Pred. No. 8.6e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLQGVGLGGPGAGSLQPLALEG 60

DB 3 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLQGVGLGGPGAGSLQPLALEG 62

QY 61 SLOKRGIVECCCTSGISLYOLENYCN 86

DB 63 SLOKRGIVECCCTSGISLYOLENYCN 88

RESULT 14

AA07687
ID AAR07687 standard; Protein: 92 AA.

XX AC AAR07687;

XX AC

XX 25-MAR-2003 (updated)

DT 21-APR-1992 (first entry)

XX XX

DE Yeast alpha-factor signal-human proinsulin fusion protein
 XX
 KW BCA-5; yeast preferred codons; post-translational processing;
 KW insulin; endopeptidase.
 XX

CS Synthetic.

FH Key Location/Qualifiers

FI Cleavage-site 6..7

FI /note- "signal-proinsulin fusion"

FI Cleavage-site 37..38

PN US5077264-A.

XX 31-DEC-1991.

XX 08-APR-1988; 88US-0184252.

XX 21-JUN-1984; 84US-0624308.

PR 08-APR-1988; 88US-0184252.

XX (REGC) UNIV CALIFORNIA.

XX Brake A7, Blair LC, Julius D, Thorne JW.

XX WPI; 1982-032671/04.

DR N-PSDB; AAQ20543.

XX Novel DNA for endopeptidase prodn. - useful for in vivo or in vitro processing of polypeptide(s)

XX Example 1; Fig 1; 16pp; English.

XX The fusion product is encoded by a synthetic sequence having at its 5'-end a modification of the 3'-end of the naturally occurring alpha-factor secretory leader and processing signal sequence, where three G-A pairs have been deleted. A plasmid containing the synthetic proinsulin coding sequence was used to transform KEX2 mutant yeast strains in the presence or absence of the cloned KEX2 gene. Post-translational processing of pro-insulin into peptides only occurred in yeast transformed to KEX2 plus. See also AAQ20545. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 92 AA:

Query Match 100.0%; Score 463; DB 13; Length 92;

Best Local Similarity 100.0%; Pred. No. 9e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVKALYLVCGRGFFYPKTRREAEDLVQGVDELGGPGAGSLQPLALEG 60

DB 7 FVNHLCGSHLVKALYLVCGRGFFYPKTRREAEDLVQGVDELGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVECCCTSGCSLYQLENYCN 86

DB 67 SLOKRGIVECCCTSGCSLYQLENYCN 92

RESULT 15

AAAP90102

ID AAAP90102 standard; protein; 93 AA.

AC AAAP90102;

XX 25-MAR-2003 (updated)

XX 01-NOV-1989 (first entry)

XX Synthetic proinsulin.

XX Proinsulin; synthetic gene; yeast.

XX Homo sapiens.

XX

PE Key Location/Qualifiers

FI Peptide 8..37

XX /note- "proinsulin chain B"

XX Region 38..39

XX /note- "dibasic processing site"

XX Peptide 38..72

XX Region 71..72

XX /note- "dibasic processing site"

XX Peptide 73..93

XX /note- "proinsulin chain A"

XX EP424274-A.

XX 13-JUL-1989.

XX 23-DEC-1986; 85EP-0312406.

XX 30-DEC-1987; 87US-0135682.

XX (CHIR) CHIRON CORP.

XX Tekampoliso P;

XX WP; 1985-208260/29.

XX N-PSDB; AAN90279.

XX Expression and secretion of heterologous proteins in yeast - using truncated a-factor leader sequence contg. signal peptide and glycosylation site.

XX Disclosure; fig 1; 22pp; English.

XX The proinsulin is expressed in Saccharomyces cerevisiae A3103.1. It is encoded by a synthetic gene (see AAN90279). The 3 consecutive peptides are proinsulin chains B, C and A resp. The regions are dibasic processing sites.

XX (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 93 AA:

Query Match 100.0%; Score 463; DB 10; Length 93;

Best Local Similarity 100.0%; Pred. No. 9.1e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVKALYLVCGRGFFYPKTRREAEDLVQGVDELGGPGAGSLQPLALEG 60

DB 8 FVNHLCGSHLVKALYLVCGRGFFYPKTRREAEDLVQGVDELGGPGAGSLQPLALEG 67

QY 61 SLOKRGIVECCCTSGCSLYQLENYCN 86

DB 68 SLOKRGIVECCCTSGCSLYQLENYCN 93

Search completed: September 15, 2003, 12:00:58

Job time : 62.4158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:57:40 : Search time 19.7276 Seconds
(without alignments)
184,449 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 FVNHLCGSHLVEALYLVCGERGFFYYTPKTRRRAEHLQVGVQLGGGSGNGSLQPLADES 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328777 seqs, 4211058 residues

Total number of hits satisfying chosen parameters: 26717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:
1: /cgn2_6/ptodata/1/iaa/5A_CGMS.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_CGMS.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_CGMS.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_CGMS.pep.*
5: /cgn2_6/ptodata/1/iaa/PC/DUS_CGMS.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	463	100.0	86	4	US-09-477-924-2
2	463	100.0	86	4	US-09-723-981-2
3	463	100.0	86	4	US-09-723-896-2
4	463	100.0	86	4	US-09-874-380-1
5	463	100.0	96	2	US-09-134-836-4
6	463	100.0	96	4	US-09-386-303A-4
7	463	100.0	97	1	US-08-160-376A-4
8	463	100.0	110	3	US-08-950-720A-11
9	463	100.0	110	3	US-08-589-025-2
10	463	100.0	110	3	US-08-784-582-2
11	463	100.0	110	3	US-08-785-271-2
12	463	100.0	110	4	US-08-472-701-2
13	463	100.0	110	4	US-09-185-852-2
14	463	100.0	110	5	PCT-US95-08596-2
15	463	100.0	117	4	US-09-260-030-63
16	463	100.0	130	4	US-09-285-030-62
17	463	100.0	151	2	US-08-508-664-15
18	463	100.0	161	2	US-08-508-664-16
19	463	100.0	167	1	US-07-918-953-8
20	463	100.0	167	1	US-08-081-661-8
21	457	98.7	96	2	US-09-134-836-5
22	457	98.7	96	4	US-09-386-303A-5
23	457	98.7	97	1	US-08-389-487-2
24	456	98.5	90	1	US-08-030-731A-43
25	446	96.3	97	3	US-09-099-307-6
26	444	95.9	97	3	US-09-099-307-8
27	443	95.7	110	3	US-08-589-025-4

```

28      443      95.7      110      3      US-08-784-582-4      Sequence 4, Appli
29      443      95.7      110      3      US-08-785-271-4      Sequence 4, Appli
30      440      95.0      97      3      US-09-099-307-7      Sequence 7, Appli
31      435      94.0      97      3      US-09-099-307-11     Sequence 11, Appli
32      398      86.0      91      4      US-09-676-787-7      Sequence 7, Appli
33      292.5      63.2      67      4      US-08-981-988A-1     Sequence 1, Appli
34      290.5      62.7      83      4      US-08-981-988A-3     Sequence 3, Appli
35      288.5      62.3      83      4      US-08-981-988A-6     Sequence 6, Appli
36      286.5      61.9      67      4      US-08-981-988A-2     Sequence 16, Appli
37      285.5      61.7      83      4      US-08-981-988A-8     Sequence 2, Appli
38      284.5      61.4      67      4      US-08-981-988A-5     Sequence 8, Appli
39      284.5      60.8      57      4      US-08-981-988A-7     Sequence 5, Appli
40      281.5      60.7      86      4      US-09-201-227A-43    Sequence 7, Appli
41      281      60.5      112     4      US-09-484-848-15     Sequence 43, Appli
42      279.5      60.4      83      4      US-08-981-988A-9     Sequence 15, Appli
43      278.5      60.2      59      4      US-08-981-988A-4     Sequence 9, Appli
44      272.5      58.9      53      4      US-08-233-617-4     Sequence 4, Appli
45

```

ALIGNMENTS

```

RESULT 1
US-09-477-924-2
: Sequence 2, Application: US/09477924
: Patent No. 6403764
: GENERAL INFORMATION:
: APPLICANT: Dubaquitte, Yves
: APPLICANT: Lowman, Henry
: TITLE OF INVENTION: PROTEIN VARIANTS
: FILE REFERENCE: P1712R1-1
: CURRENT APPLICATION NUMBER: US/09/477,924
: CURRENT FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 2
: LENGTH: 86
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-477-924-2

```

```

Query Match      100.0%   Score 463:   DB 4:   Length 86;
Best Local Similarity 100.0%   Pred. No. 1.5e-47;
Matches      86:   Conservative      0:   Mismatches      0:   Indels      0:   Gaps      0:

```

```

CY      1      FVNHLCGSHLVEALYLVCGERGFFYYTPKTRRRAEHLQVGVQLGGGSGNGSLQPLADES 50
ID      1      FVNHLCGSHLVEALYLVCGERGFFYYTPKTRRRAEHLQVGVQLGGGSGNGSLQPLADES 50
QY      61      SLCKRGIVECCCTSCISLYOLENYCN 86
DB      61      SLCKRGIVECCCTSCISLYOLENYCN 86

```

```

RESULT 2
US-09-723-981-2
: Sequence 2, Application US/09723981
: Patent No. 6508874
: GENERAL INFORMATION:
: APPLICANT: Dubaquitte, Yves
: APPLICANT: Lowman, Henry
: TITLE OF INVENTION: PROTEIN VARIANTS
: FILE REFERENCE: P1712R1
: CURRENT APPLICATION NUMBER: US/09/723,981
: CURRENT FILING DATE: 2000-11-28
: PRIOR APPLICATION NUMBER: 09/477,923
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 2
: LENGTH: 86
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-723-981-2

```



```
US-09-386-303A-4
: Sequence 4, Application US/09/36303A
: Patent No. 6380355
: GENERAL INFORMATION:
: APPLICANT: Rubroder, Franz-Josef
: TITLE OF INVENTION: Improved process for obtaining
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farrelow, Garrett &
: Dunner
: STREET: 1300 7 Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3415
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/386,303A
: FILING DATE: 31-Aug-1999
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/134,846
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Leslie McDermott
: REGISTRATION NUMBER: 34,872
: REFERENCE/DOCKET NUMBER: 02481.1609-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4500
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 96 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: FEATURES:
: NAME/KEY: Protein
: LOCATION: 1-96
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-386-303A-4

Query Match 100.0%; Score 463; DB 4; Length 96;
Best Local Similarity 100.0%; Prod. No. 1.8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNHQGGSHLVKALYVCGERGFFYTPKTRREAEQLQGVGVGGSGGAGSLQPLAERG 60
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Db 11 FVNHQGGSHLVKALYVCGERGFFYTPKTRREAEQLQGVGVGGSGGAGSLQPLAERG 70
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Qy 61 SLOKRGIVECCCTSCSLYQLENYCN 86
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Db 71 SLOKRGIVECCCTSCSLYQLENYCN 96
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+

RESULT 7
US-08-160-376A-4
: Sequence 4, Application US/08/160376A
: Patent No. 5473049
: GENERAL INFORMATION:
: APPLICANT: Obermeyer, Ranier
: APPLICANT: Gerl, Martin
: APPLICANT: Ludwig, Jurgen
: APPLICANT: Sabel, Walter
```

```
US-08-160-376A-4
: Sequence 4, Application US/08/160376A
: Patent No. 5473049
: GENERAL INFORMATION:
: APPLICANT: Obermeyer, Ranier
: APPLICANT: Gerl, Martin
: APPLICANT: Ludwig, Jurgen
: APPLICANT: Sabel, Walter
: TITLE OF INVENTION: Process For Obtaining Prinsalin
: TITLE OF INVENTION: Possessing Correctly Linked
: TITLE OF INVENTION: Cystine Bridges
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenneth A. Genoni, Esq.
: STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
: CITY: Somerville
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 08876-1258
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
: COMPUTER: IBM 386
: OPERATING SYSTEM: WINDOWS 3.1
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/160,376A
: FILING DATE: December 1, 1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GE P 4240420.7
: FILING DATE: December 2, 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Barbara V. Maurer, Esq.
: REGISTRATION NUMBER: 31,287
: REFERENCE/DOCKET NUMBER: HCE 92/P 384
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 231-4079
: TELEFAX: (908) 231-2255
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 97 Amino Acids
: TYPE: Amino Acid (AA)
: TOPOLOGY: not relevant
US-08-160-376A-4

Query Match 100.0%; Score 463; DB 1; Length 97;
Best Local Similarity 100.0%; Prod. No. 1.8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNHQGGSHLVKALYVCGERGFFYTPKTRREAEQLQGVGVGGSGGAGSLQPLAERG 60
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Db 12 FVNHQGGSHLVKALYVCGERGFFYTPKTRREAEQLQGVGVGGSGGAGSLQPLAERG 71
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Qy 61 SLOKRGIVECCCTSCSLYQLENYCN 86
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+
Db 72 SLOKRGIVECCCTSCSLYQLENYCN 97
+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+|||||+

RESULT 8
US-08-950-720A-11
: Sequence 11, Application US/89950720A
: Patent No. 6046028
: GENERAL INFORMATION:
: APPLICANT: Conklin, Darrell C.
: APPLICANT: Loton-Day, Catherine E.
: APPLICANT: Lok, Si
: APPLICANT: Jaspers, Stephen R.
: TITLE OF INVENTION: INSULIN HOMOLOG
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ZymoGenetics, Inc.
: STREET: 1201 Eastlake Avenue East
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/03/950,720A
3 FILING DATE:
4 CLASSIFICATION: 435
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER:
7 FILING DATE:
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Sawislak, Deborah A.
10 REGISTRATION NUMBER: 37,438
11 REFERENCE/DOCKET NUMBER: 95-09
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 206-442-6672
14 TELEFAX: 206-442-6678
15
16 INFORMATION FOR SEQ ID NO: 1:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 110 amino acids
19 TYPE: amino acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 MOLECULE TYPE: No. 60460284
23
24 US-08-950-720A-11

Query Match 100.0% Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNCHLGGSHLVEALYVCGSGFFYTPKTRAEALDVGQVGLGGPGAGSLQPLALEG 60
DB 25 FVNCHLGGSHLVEALYVCGSGFFYTPKTRAEALDVGQVGLGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEQCCTSCSLYQLENYCN 86
DB 85 SLOKRGIVEQCCTSCSLYQLENYCN 110

RESULT 9

US-08-589-028-2
1 Sequence 2, Application US/08519028
2 Patent No. 6087129
3
4 GENERAL INFORMATION:
5 APPLICANT: Newgard, Christopher B.
6 APPLICANT: Haiban, Philippe A.
7 APPLICANT: No. 6087129minington, Karl D.
8 APPLICANT: Clark, Samuel A.
9 APPLICANT: Thigpen, Anice E.
10 APPLICANT: Quade, Christian
11 APPLICANT: Kruse, Fred
12
13 TITLE OF INVENTION: Recombinant Expression of Proteins from
14 TITLE OF INVENTION: SECRETORY CELL LINES
15
16 NUMBER OF SEQUENCES: 50
17 CORRESPONDENCE ADDRESS:
18 ADDRESSEE: Arnold, White & Durkee
19 STREET: P. O. Box 4433
20 CITY: Houston
21 STATE: TX
22 COUNTRY: USA
23 ZIP: 77210-4433
24
25 COMPUTER READABLE FORM:
26 MEDIUM TYPE: Floppy disk
27 COMPUTER: IBM PC compatible
28 OPERATING SYSTEM: PC-DOS/MS-DOS
29 SOFTWARE: PatentIn Release #1.0, Version #1.30
30
31 CURRENT APPLICATION DATA:
32 APPLICATION NUMBER: 05/09/589,028
33 FILING DATE: Concurrently Herewith
34 CLASSIFICATION: 435
35
36 PRIOR APPLICATION DATA:
37 APPLICATION NUMBER: US 08/589,028
38 FILING DATE: 19-JAN-1996
39 ATTORNEY/AGENT INFORMATION:
40 NAME: Highland, Steven L.
41 REGISTRATION NUMBER: 37,642
42 REFERENCE/DOCKET NUMBER: UTSD:514
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE: 512/418-3000
45 TELEFAX: 512/474-7577
46
47 INFORMATION FOR SEQ ID NO: 2:
48 SEQUENCE CHARACTERISTICS:
49 LENGTH: 110 amino acids
50 TYPE: amino acid
51 STRANDEDNESS:

1 TELEPHONE: (512) 418-3000
2 TELEFAX: (512) 474-7577
3 INFORMATION FOR SEQ ID NO: 2:
4 SEQUENCE CHARACTERISTICS:
5 LENGTH: 110 amino acids
6 TYPE: amino acid
7 STRANDEDNESS:
8 TOPOLOGY: linear
9 US-08-589-028-2
10
11 Query Match 100.0% Score 464; DB 3; Length 110;
12 Best Local Similarity 100.0%; Pred. No. 2.1e-47;
13 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
14
15 QY 1 FVNCHLGGSHLVEALYVCGSGFFYTPKTRAEALDVGQVGLGGPGAGSLQPLALEG 60
16 DB 25 FVNCHLGGSHLVEALYVCGSGFFYTPKTRAEALDVGQVGLGGPGAGSLQPLALEG 60
17 QY 61 SLOKRGIVEQCCTSCSLYQLENYCN 86
18 DB 85 SLOKRGIVEQCCTSCSLYQLENYCN 110

RESULT 10

US-08-784-592-2
1 Sequence 2, Application US/98784582
2 Patent No. 6110707
3
4 GENERAL INFORMATION:
5 APPLICANT: Newgard, Christopher B.
6 APPLICANT: Haiban, Philippe A.
7 APPLICANT: No. 6110707minington, Karl D.
8 APPLICANT: Clark, Samuel A.
9 APPLICANT: Thigpen, Anice E.
10 APPLICANT: Quade, Christian
11 APPLICANT: Kruse, Fred
12 APPLICANT: McGarity, Dennis
13
14 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
15 TITLE OF INVENTION: SECRETORY CELL LINES
16
17 NUMBER OF SEQUENCES: 79
18 CORRESPONDENCE ADDRESS:
19 ADDRESSEE: Arnold, White & Durkee
20 STREET: P. O. Box 4433
21 CITY: Houston
22 STATE: Texas
23 COUNTRY: USA
24 ZIP: 77210
25
26 COMPUTER READABLE FORM:
27 MEDIUM TYPE: Floppy disk
28 COMPUTER: IBM PC compatible
29 OPERATING SYSTEM: PC-DOS/MS-DOS
30 SOFTWARE: PatentIn Release #1.0, Version #1.30
31
32 CURRENT APPLICATION DATA:
33 APPLICATION NUMBER: US/08/784,582
34 FILING DATE: Concurrently Herewith
35 CLASSIFICATION: 435
36
37 PRIOR APPLICATION DATA:
38 APPLICATION NUMBER: US 60/028,427
39 FILING DATE: 15-OCT-1996
40
41 APPLICATION NUMBER: US 08/589,028
42 FILING DATE: 19-JAN-1996
43 ATTORNEY/AGENT INFORMATION:
44 NAME: Highland, Steven L.
45 REGISTRATION NUMBER: 37,642
46 REFERENCE/DOCKET NUMBER: UTSD:514
47 TELECOMMUNICATION INFORMATION:
48 TELEPHONE: 512/418-3000
49 TELEFAX: 512/474-7577
50
51 INFORMATION FOR SEQ ID NO: 2:
52 SEQUENCE CHARACTERISTICS:
53 LENGTH: 110 amino acids
54 TYPE: amino acid
55 STRANDEDNESS:

TOPOLOGY: linear
US-08-784-582-2

Query Match
Best Local Similarity 100.0%; Score 463; DB 3; Length 110;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVECCCTCSLSLYQLENYCN 86
DB 85 SLQKRGIVECCCTCSLSLYQLENYCN 110

RESULT 11
US-08-785-271-2
; Sequence 2, Application US/08785271
; Patent No. 6194176
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6194176minetox, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Arico E.
; APPLICANT: Cuevas, Christian
; APPLICANT: Kruse, Fred
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PRE-IL1NS FROM
; TITLE OF INVENTION: SECRETORY CELL LINES
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Barker
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,271
; FILING DATE: Conventionally Related
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-785-271-2

Query Match
Best Local Similarity 100.0%; Score 463; DB 3; Length 110;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVECCCTCSLSLYQLENYCN 86
DB 85 SLQKRGIVECCCTCSLSLYQLENYCN 110

RESULT 12
US-08-472-701-2
; Sequence 2, Application US/08472701
; Patent No. 6509165
; GENERAL INFORMATION:
; APPLICANT: Griffin, Ann C.
; APPLICANT: Hickey, William F.
; TITLE OF INVENTION: Detection and Treatment Methods for
; TITLE OF INVENTION: Type 1 Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Gallo A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-472-701-2

Query Match
Best Local Similarity 100.0%; Score 463; DB 4; Length 110;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVECCCTCSLSLYQLENYCN 86
DB 85 SLQKRGIVECCCTCSLSLYQLENYCN 110

RESULT 13
US-09-185-852-2
; Sequence 2, Application US/09185852
; Patent No. 6537806
; GENERAL INFORMATION:
; APPLICANT: Osborne, William R.A.
; APPLICANT: Ramesh, Nagarajan
; TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
; FILE REFERENCE: P-0W 3264
; CURRENT APPLICATION NUMBER: US/09/185,852

```
? CURRENT FILING DATE: 1998-11-04
? EARLIER APPLICATION NUMBER: 60/087,660
? EARLIER FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 1
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 2
? LENGTH: 110
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-185-852-2

Query Match      100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 60
DB 25 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 86
61 SLOKRGIVEQCCTSGSLYOLENYCN 86
85 SLOKRGIVEQCCTSGSLYOLENYCN 110

RESULT 15
US-09-280-030-63
? Sequence 63, Application: US/99280030A
? Patent No. 6506595
? GENERAL INFORMATION:
? APPLICANT: Sato, Seiji
? APPLICANT: Higashikuni, Naohiko
? APPLICANT: Kudo, Toshiyuki
? APPLICANT: Kondo, Masaki
? TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
? TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
? FILE REFERENCE: DNAS
? FILE REFERENCE: 382,1026
? CURRENT APPLICATION NUMBER: US/99/280,030A
? EARLIER FILING DATE: 1999-03-26
? EARLIER FILING DATE: 1999-03-31
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 63
? LENGTH: 117
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Designated is
? OTHER INFORMATION: an amino acid sequence of
? OTHER INFORMATION: MW2sp-MW2pml0-Met-Proinsulin
US-09-280-030-63

Query Match      100.0%; Score 463; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2,2e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 60
DB 32 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 91
61 SLOKRGIVEQCCTSGSLYOLENYCN 86
92 SLOKRGIVEQCCTSGSLYOLENYCN 117

Search completed: September 15, 2003, 12:05:32
Job time : 20.7276 secs

? CURRENT FILING DATE: 1998-11-04
? EARLIER APPLICATION NUMBER: 60/087,660
? EARLIER FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 1
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 2
? LENGTH: 110
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-185-852-2

Query Match      100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 60
DB 25 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 86
61 SLOKRGIVEQCCTSGSLYOLENYCN 86
85 SLOKRGIVEQCCTSGSLYOLENYCN 110

RESULT 14
PCT-US95-08596-2
? Sequence 2, Application: PCT/US9508596
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: Proinsulin Peptide Compounds for Diabetes
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: LAHIVE & COCKFIELD
? STREET: 50 State Street, Suite 510
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02109-1875
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII Text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/08596
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/272,220
? FILING DATE: 08-JULY-1993
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: DeConti, Giulio A., Jr.
? REGISTRATION NUMBER: 31,103
? REFERENCE/DOCKET NUMBER: PCT-952PC
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617)227-7400
? TELEFAX: (617)227-5942
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 110 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
PCT-US95-08596-2

Query Match      100.0%; Score 463; DB 5; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 60
DB 32 FVNOHLCGSHLVALYVCGERFFYTPKTRREAEDLVQVQLGGPGAGSLQPLALEG 91
61 SLOKRGIVEQCCTSGSLYOLENYCN 86
92 SLOKRGIVEQCCTSGSLYOLENYCN 117

Search completed: September 15, 2003, 12:05:32
Job time : 20.7276 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 12:03:35 : Search time 36.0645 seconds
(without alignments)
347,945 Million cell updates/sec

Title: US-09-423-100-4

Perfect score: 463

Sequence: 1 FVNHLCGSHLVEALYLVCGERGFFYYTPKTRREAEDLQVQLGGPGAGSLOPLALEG 60

Scoring table: BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pept.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pept.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pept.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pept.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pept.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pept.*
7: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pept.*
8: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pept.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pept.*
10: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pept.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pept.*
12: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pept.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pept.*
14: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pept.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pept.*
16: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pept.*
17: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pept.*
18: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUBCOMB.pept.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	86	10	US-09-878-380-1
2	463	100.0	86	11	US-09-858-935B-4
3	463	100.0	86	14	US-10-028-410-2
4	463	100.0	86	14	US-10-054-873-4
5	463	100.0	96	10	US-09-947-563-4
6	463	100.0	110	9	US-09-205-658-125
7	463	100.0	110	9	US-09-815-229-3
8	463	100.0	110	10	US-09-804-409A-9
9	463	100.0	110	12	US-09-969-748C-6
10	463	100.0	110	15	US-10-038-686-1
11	463	100.0	110	15	US-10-328-813-2
12	463	100.0	117	9	US-09-280-030-63
13	463	100.0	130	9	US-09-280-030-62
14	457	98.7	96	10	US-09-947-563-5
15	306	66.1	166	9	US-09-925-257-805

19	300	64.8	56	9	US-09-815-229-10	Sequence 10, Appl
20	285	61.6	54	9	US-09-815-229-13	Sequence 13, Appl
21	267	57.7	52	14	US-10-054-873-5	Sequence 5, Appl
22	267	57.7	107	14	US-10-054-873-6	Sequence 6, Appl
23	267	57.7	150	14	US-10-054-873-7	Sequence 7, Appl
24	261.5	56.5	51	11	US-09-858-935B-5	Sequence 5, Appl
25	261.5	56.5	51	14	US-10-028-410-3	Sequence 3, Appl
26	259.5	55.8	124	10	US-09-804-711-12	Sequence 16, Appl
27	259.5	55.8	138	10	US-09-861-687-19	Sequence 18, Appl
28	256	55.3	50	14	US-10-066-009A-3	Sequence 19, Appl
29	250.5	54.1	124	9	US-09-736-611-12	Sequence 3, Appl
30	250.5	54.1	125	9	US-09-736-611-10	Sequence 12, Appl
31	249.5	53.9	147	9	US-09-736-611-8	Sequence 10, Appl
32	249.5	53.9	124	9	US-09-740-359-12	Sequence 8, Appl
33	249.5	53.9	124	10	US-09-894-711-12	Sequence 12, Appl
34	248.5	53.9	125	9	US-09-740-359-10	Sequence 10, Appl
35	248.5	53.9	125	10	US-09-894-711-10	Sequence 10, Appl
36	248	53.6	144	9	US-09-740-359-7	Sequence 7, Appl
37	248	53.6	144	9	US-09-736-611-6	Sequence 6, Appl
38	248	53.6	146	10	US-09-740-359-5	Sequence 5, Appl
39	227.5	49.1	144	10	US-09-894-711-5	Sequence 5, Appl
40	209	45.1	144	10	US-09-894-711-7	Sequence 7, Appl
41	197	42.5	46	9	US-09-205-658-132	Sequence 132, App
42	184.5	39.8	50	9	US-09-205-658-133	Sequence 133, App
43	180	38.9	46	9	US-09-205-658-115	Sequence 115, App
44	178	38.4	32	9	US-09-205-658-135	Sequence 135, App
45	178	38.4	32	9	US-09-815-229-14	Sequence 14, Appl
46	175	37.8	46	10	US-09-947-563-6	Sequence 6, Appl
47	174	37.6	35	9	US-09-205-658-134	Sequence 134, App
48					US-09-815-229-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-878-380-1
Sequence 1, Application US/09678380
Patent No. US20020160435A1
GENERAL INFORMATION:
APPLICANT: Fujirebio Inc.
APPLICANT: KITAJIMA, Sachiko
APPLICANT: KURANO, Yoshihiro
APPLICANT: NAKATSUBO, Kaoru
APPLICANT: NISHIZONO, Isao
TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
FILE REFERENCE: C760-0291P
CURRENT APPLICATION NUMBER: US/09/878,380
PRIOR APPLICATION NUMBER: 2001-06-12
PRIOR FILING DATE: 2000-06-12
NUMBER OF SEQ IDS NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 86
TYPE: PPT
ORGANISM: Homo sapiens
US-09-878-380-1

Query Match 100.0% Score 463, DB ID: Length 86,
Best local Similarity 100.0% Pred. No. 9, 7e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q1 1 FVNHLCGSHLVEALYLVCGERGFFYYTPKTRREAEDLQVQLGGPGAGSLOPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYYTPKTRREAEDLQVQLGGPGAGSLOPLALEG 60

Q7 61 SLQKRGIVEQCCTSCISLYOLENYCN 86

DB 62 SLQKRGIVEQCCTSCISLYOLENYCN 86

RESULT 2

7

Db 85 SLOKRGIVECCCTSCSYOLENYCN 110

RESULT 9

US-09-969-748C-6
Sequence 6, Application US/5969748C
Publication No. US20030161629A1
GENERAL INFORMATION:
APPLICANT: ARZEXE PHARMACEUTICALS, INC.
APPLICANT: HOUSTON, Lou, L.
APPLICANT: SHERIDAN, Philip, J.
APPLICANT: HAWLEY, Stephen
APPLICANT: GYNN, Jacqueline, M.
APPLICANT: CHAPIN, Steven
APPLICANT: BASU, Amaresh
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF BLOOD-CELLS ACTIVELY
FILE REFERENCE: US/5720-0303
CURRENT FILING DATE: 2002-12-10
PRIOR FILING DATE: 2001-02-04
PRIOR APPLICATION NUMBER: US 56/248,819
PRIOR FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: US 60/248,478
PRIOR FILING DATE: 2000-11-11
PRIOR APPLICATION NUMBER: US 60/247,929
PRIOR FILING DATE: 2000-10-07
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-748C-6

Query Match 100.0%; Score 463; DB 12; Length 110;

Best Local Similarity 100.0%; Pred. No. 1.3e-44;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 60

DB 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVECCCTSCSYOLENYCN 86

DB 85 SLOKRGIVECCCTSCSYOLENYCN 110

RESULT 10

US-10-038-686-1
Sequence 1, Application US/10038686
Publication No. US20030045467A1
GENERAL INFORMATION:
APPLICANT: Otsuka, Tibeter
TITLE OF INVENTION: ANTIGEN VACCINE
FILE REFERENCE: 10276-067001
CURRENT APPLICATION NUMBER: US/10/038,586
CURRENT FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/269,062
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq 1.0; Windows Version 4.0
SEQ ID NO 1
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-686-1

Query Match 100.0%; Score 463; DB 15; Length 110;

Best Local Similarity 100.0%; Pred. No. 1.3e-44;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 60

DB 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVECCCTSCSYOLENYCN 86

DB 85 SLOKRGIVECCCTSCSYOLENYCN 110

RESULT 11

US-10-328-813-2
Sequence 2, Application US/10328813
Publication No. US20030113305A1
GENERAL INFORMATION:
APPLICANT: Osborne, William R.A.
APPLICANT: Ramesh, Nagarajan
TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
FILE REFERENCE: P-CW 3264
CURRENT APPLICATION NUMBER: US/10/328,813
CURRENT FILING DATE: 2002-12-23
PRIOR APPLICATION NUMBER: US/09/185,852
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: 60/087,660
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-328-813-2

Query Match 100.0%; Score 463; DB 15; Length 110;

Best Local Similarity 100.0%; Pred. No. 1.3e-44;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 60

DB 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQGVLEGGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVECCCTSCSYOLENYCN 86

DB 85 SLOKRGIVECCCTSCSYOLENYCN 110

RESULT 12

US-09-280-030-63
Sequence 63, Application US/09280030A
Patent No. US20010021515A1
GENERAL INFORMATION:
APPLICANT: Sato, Seiji
APPLICANT: Higashikuni, Nachiko
APPLICANT: Kudo, Toshiyuki
APPLICANT: Kondo, Masaaki
TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
FILE REFERENCE: DNAS
FILE REFERENCE: 382.1026
CURRENT APPLICATION NUMBER: US/09/280,030A
CURRENT FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: JP10-87339/1998
EARLIER FILING DATE: 1998-03-31
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 63
LENGTH: 117
TYPE: PRT
ORGANISM: Artificial Sequence

Query Match 100.0%; Score 463; DB 15; Length 110;

Best Local Similarity 100.0%; Pred. No. 1.3e-44;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 SLOKRGIVECCCTSCSYOLENYCN 86

DB 85 SLOKRGIVECCCTSCSYOLENYCN 110

```

Query Match      100.0% Score 463; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 86; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 32 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 61

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 86
DB 92 SLOKRGIVEQCCTSCSLYQLENYC 117

RESULT 13
US-09-280-030-62
Sequence 62, Application US/09280030A
Patent No. US20030021513A1
GENERAL INFORMATION:
APPLICANT: Satg, Selji
APPLICANT: Higashikuni, Naohiko
APPLICANT: Kudo, Toshiyuki
APPLICANT: Kondo, Masaaki
TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
FILE REFERENCE: 382,1026
CURRENT APPLICATION NUMBER: US/99/285,030A
EARLIER FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: JP10-87439/-99#
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 62
LENGTH: 130
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Postulated is
OTHER INFORMATION: an amino acid sequence of
OTHER INFORMATION: MWSP:MWp10-(His)6-Linker-Met-Proinsulin
US-09-280-030-62

Query Match      100.0% Score 463; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 86; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 45 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 164

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 86
DB 105 SLOKRGIVEQCCTSCSLYQLENYC 130

RESULT 14
US-09-947-563-5
Sequence 5, Application US/9947563
Patent No. US20030150234A1
GENERAL INFORMATION:
APPLICANT: Rubtoder, Franz-Josef
APPLICANT: Keller, Reinhold
TITLE OF INVENTION: Improved process for obtaining
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Finnegan, Henderson, Farabee, Garrett &
Derner
STREET: 1300 I Street, N.W.
City: Washington
STATE: D.C.
COUNTRY: USA

```

```

ZIP: 20005-3325
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/947,563
FILING DATE: 07-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,836
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Leslie McDonnell
REGISTRATION NUMBER: 34,972
REFERENCE/DOCKET NUMBER: 02481.1600-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Protein
LOCATION: 1..96
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-947-563-5

Query Match      98.7% Score 457; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 5.2e-44;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 11 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 70

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 85
DB 71 SLOKRGIVEQCCTSCSLYQLENYC 95

RESULT 15
US-09-925-277-805
Sequence 805, Application US/99925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/03/925,297
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: PCT/US00/05489
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 805
LENGTH: 166
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (84)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE

```


Score version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 / Search time 14,1792 seconds
(without alignment)
533,264 Million cell updates/sec

Title: US-09-423-100-4

Perfect score: 453

Sequence: 1 FVNHLCGSHLVEALYVQG.....IVFGCCSTSCSLYLENVON 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96158682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

pir_761*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	453	100.0	110	1	IPRH
2	453	100.0	110	2	A42179
3	456	98.5	110	2	A42179
4	456	98.5	110	2	JOE178
5	424	91.6	110	1	INSH
6	417	90.1	110	1	IPDG
7	394	85.1	86	1	IPRG
8	394	85.1	110	1	IPRT2
9	394	85.1	110	1	INXS2
10	392	84.7	108	2	A39843
11	392	84.7	110	2	A48156
12	385	83.2	110	1	IPRT1
13	383	82.7	84	1	IPRG
14	366.5	79.2	105	1	IPRG
15	366	79.0	108	1	INXS1
16	334.5	72.2	108	2	S92274
17	320.5	69.2	77	1	INSH
18	314	67.8	110	1	INPG
19	277.5	59.9	109	1	IPK100
20	276.5	59.7	103	2	I51221
21	265.5	57.3	106	1	IPX12
22	265.5	57.3	107	1	IPCH
23	262.5	56.7	106	1	IPX11
24	256.5	55.4	51	1	IPWHP
25	256.5	55.4	51	1	INMHF
26	256.5	55.4	51	1	INEL
27	256.5	55.4	81	1	IPDK
28	256	55.3	96	2	PC7082
29	254.5	55.0	51	1	INRY

ALIGNMENTS

RESULT 1

IPRH

Insulin precursor [validated] - human

N:Alternate names: preproinsulin

C:Species: Homo sapiens (man)

C:Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text change 08-Dec-2000

C:Accession: A93222; A94253; A93216; A94251; A93144; A92075; A9186; I58114; A0157

R:Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.

Nature 284, 26-32, 1980

A:Title: Sequence of the human insulin gene.

A:Reference number: A93222; MUID:80120725; PMID:6243748

A:Accession: A93222

A:Molecule type: DNA

A:Residues: 1-110 <BELL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Ullrich, A.; Dulli, T.J.; Gray, A.; Brosius, J.; Sures, I.

Science 205, 612-615, 1980

A:Title: Genetic variation in the human insulin gene.

A:Reference number: A94253; MUID:80236313; PMID:6248962

A:Accession: A94253

A:Molecule type: DNA

A:Residues: 1-110 <ULL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Ullrich, A.; Dulli, T.J.; Gray, A.; Brosius, J.; Cordell, B.; Goodman, H.M.; Rutter, W.J.

Nature 282, 525-527, 1979

A:Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.

A:Reference number: A93216; MUID:80054779; PMID:501234

A:Accession: A93216

A:Molecule type: mRNA

A:Residues: 1-110 <SEEL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Sures, I.; Soedel, D.V.; Gray, A.; Ullrich, A.

Science 205, 57-59, 1980

A:Title: Nucleotide sequence of human preproinsulin complementary DNA.

A:Reference number: A94251; MUID:80147417; PMID:6927840

A:Accession: A94251

A:Molecule type: mRNA

A:Residues: 1-110 <SURE>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Nicol, D.S.H.W.; Smith, I.F.

Nature 187, 483-485, 1960

A:Title: Amino-acid sequence of human insulin.

A:Reference number: A93144

A:Accession: A93144

A:Molecule type: protein

A:Residues: 25-54; 90-110 <NIC>

R:Over, P.F.; Cho, S.; Peterson, J.D.; Steiner, D.F.

J. Biol. Chem. 246, 1375-1386, 1971

A:Title: Studies on human proinsulin. Isolation and amino acid sequence of the human proinsulin.

A:Reference number: A92075; MUID:71116410; PMID:5101771

A:Accession: A92075

A:Molecule type: protein

A:Residues: 57-87 <OYE>
 R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20: 190-199, 1971
 A:Title: Amino acid sequence of the C-peptide of human proinsulin.
 A:Reference number: A91196; MUII:71257722; PMID:560484
 A:Accession: A91186
 A:Molecule type: protein.
 A:Residues: 57-87 <KOA>
 R:Lucassen, A.M.; Julier, C.; Bressi, J.P.; Roitard, C.; Froquet, F.; Lathrop, M.; Bell
 Nature Genet. 4: 305-310, 1993
 A:Title: Susceptibility to insulin dependent diabetes mellitus maps to a 5.1 kb segment
 A:Reference number: I58114; MUII:93454428; PMID:8456440
 A:Accession: I58114
 A:Status: preliminary; translated from GB/EMBL/DBE
 A:Molecule type: DNA
 A:Residues: 1-59, 63-110 <SES>
 A:Cross-references: GB:U5440; NID:9307571; PIDN:CAA43403.1; PID:938252
 R:Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Knicker, R.; Ritzel, W.
 Helv. Chim. Acta 57: 2617-2621, 1974
 A:Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbrücken.
 A:Reference number: A91636; MUII:75072277; PMID:444399
 A:Contents: annotation; synthesis
 A:Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identical
 A:Note: article in German with English abstract
 R:Naithani, V.K.
 Hoppe-Seyler's Z. Physiol. Chem. 354: 459-472, 1974
 A:Title: The synthesis of C-peptide of human proinsulin.
 A:Reference number: A91638; MUII:75040071; PMID:443504
 A:Contents: annotation; synthesis of residues 57-87
 R:Geiger, R.; Jaquet, G.; Koobig, W.
 Chem. Ber. 106: 2347-2352, 1973
 A:Title: Synthesis of the complete sequence of human proinsulin C-peptide and its junction
 A:Contents: annotation; synthesis
 A:Reference number: A90914
 R:Kaufmann, J.E.; Irzinger, J.G.; Halban, P.A.
 Biochem. J. 310: 869-874, 1995
 A:Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junction
 A:Reference number: S58661; MUII:96313185; PMID:7575420
 A:Contents: annotation; site-directed mutagenesis study of proteolytic processing
 C:Genetics:
 A:Gene: GDB:INS
 A:Cross-references: GDB:119349; OMIM:176730
 A:Map position: 11p15.3-1p15.5
 A:Introns: 63/1
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BCH>
 F:55-90-110/Product: insulin #status experimental <MAT>
 F:57-87/Domain: connecting C peptide #status experimental <CFEP>
 F:90-110/Domain: insulin chain A #status experimental <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 2.8e-43; Indels 0; Gaps 0;
 Matches 86; Conservative 0; Mismatches 0

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
 DQ 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 84

RESULT 2
 A42179
 Insulin precursor - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1994
 C:Accession: A42179; S22058
 R:Seino, S.; Bell, G.I.; Li, W.H.

Query Match 98.5%; Score 456; DB 2; Length 110;
 Best Local Similarity 98.8%; Pred. No. 1.6e-42; Indels 1; Gaps 0;
 Matches 85; Conservative 0; Mismatches 1

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
 DQ 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 84

RESULT 4
 A42179
 Insulin precursor - green monkey
 C:Species: Cercopithecus achlopus (green monkey, grivet)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1994
 C:Accession: A42179; A05232; S16494; S22056
 R:Seino, S.; Bell, G.I.; Li, W.H.
 Mol. Biol. Evol. 9: 193-203, 1992
 A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate
 A:Reference number: A42179; MUII:92219553; PMID:1560757
 A:Accession: A42179
 A:Molecule type: DNA
 A:Residues: 1-110 <SEI>
 A:Cross-references: EMBL:X61092; NID:922808; PIDN:CAA43405.1; PID:922809
 A:Note: sequences extracted from NCHI backbone (NCBIN:95185, NCBI:P:95194)
 R:Peterson, J.D.; Nehrlich, S.; Oyer, P.E.; Steiner, D.F.
 J. Biol. Chem. 247: 4866-4871, 1972
 A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog pro
 A:Reference number: A92111; MUII:72258016; PMID:4626369
 A:Accession: A05232
 A:Molecule type: protein
 A:Residues: 57-87 <PEP>
 C:Genetics:
 A:Introns: 63/1
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status predicted <BCH>
 F:55-90-110/Product: insulin #status predicted <MAT>
 F:57-87/Domain: connecting peptide #status experimental <CFEP>
 F:90-110/Domain: insulin chain A #status predicted <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.8%; Score 456; DB 2; Length 110;
 Best Local Similarity 98.8%; Pred. No. 1.6e-42; Indels 1; Gaps 0;
 Matches 85; Conservative 0; Mismatches 0

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
 DQ 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 84

RESULT 2
 A42179
 Insulin precursor - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1994
 C:Accession: A42179; S22058
 R:Seino, S.; Bell, G.I.; Li, W.H.

JQ0178

insulin precursor - crab-eating macaque
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:Date: 07-Sep-1990 #sequence_revision 37-Sep-1990 #text_change 16-Jul-1990
 C:Accession: JQ0178
 R:Wetzel, W.; Gronow, J.; Leineweber, M.; Wentz, M.; F. Winkler, F.
 Gene 19, 179-183, 1992
 A:Title: The nucleotide sequence of cDNA coding for proinsulin from the primate Macaca fascicularis
 A:Reference number: JQ0178; MUID:8308474; PMID:5184282

A:Accession: JQ0178
 A:Molecule type: mRNA
 A:Residues: 1-110 <MET>
 A:Cross-references: GB:J00346; NID:9342121; PIDN:AAA6549.1; PMID:842122
 C:Superfamily: insulin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin #status predicted <MAT>
 F:55-89/Domain: insulin chain B #status predicted <ACH>
 F:90-110/Domain: insulin connecting C peptide #status predicted <CPEP>
 F:91-96/Domain: insulin chain A #status predicted <ACH>
 F:91-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5% Score 456; DB 2; Length 110;

Best Local Similarity 98.8% Pred. No. 1.6e-42; Indels 0; Gaps 0;

Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

RESULT 5

INSR
 insulin precursor - rabbit
 K:Alternate names: preproinsulin
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 24-Apr-1984 #sequence_revision 25-Aug-1997 #text_change 18-Jun-1999
 C:Accession: A53438; A01581
 R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.F.; Menon, R.K.; Zahm, D.
 J. Biol. Chem. 269, 8445-8454, 1994
 A:Title: Insulin gene expression and insulin synthesis in mammalian neurons: implications for the pathogenesis of diabetes mellitus
 A:Reference number: A53438; MUID:94170230; PMID:814259
 A:Accession: A53438
 A:Molecule type: mRNA
 A:Residues: 1-110 <DEV>
 A:Cross-references: GB:003610; NID:9407970; PIDN:AAA19033.1; PMID:9407971
 R:Smith, L.F.

Am. J. Med. 40, 662-666, 1966
 A:Title: Species variation in the amino acid sequence of insulin.
 A:Reference number: A90029; MUID:65160119; PMID:5949593
 A:Accession: A01581
 A:Molecule type: protein
 A:Residues: 25-54; 90-110 <SMI>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BCH>
 F:55-89/Domain: insulin chain A #status experimental <MAT>
 F:90-110/Domain: insulin connecting C peptide #status predicted <CPEP>
 F:91-96/Domain: insulin chain A #status experimental <ACH>
 F:91-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 91.6% Score 424; DB 1; Length 110;

Best Local Similarity 90.7% Pred. No. 5e-39;

Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 ALQKRGIVEQCCTISCSLYOLENYCN 110

RESULT 6

IPNG
 insulin precursor - dog
 C:Species: Canis lupus familiaris (dog)
 C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
 C:Accession: A92413; A01587; S16493
 R:Kawox, S.C.M.; Chan, S.J.; Steiner, D.F.
 J. Biol. Chem. 258, 2357-2363, 1983
 A:Title: Cloning and nucleotide sequence analysis of the dog insulin gene. Coded at the 5' end of the insulin gene.
 A:Reference number: A92413; MUID:83109071; PMID:6296142
 A:Accession: A92413
 A:Molecule type: DNA
 A:Residues: 1-110 <SMI>
 A:Cross-references: GB:V00179; GB:J00042; NID:9994; PIDN:CAA23475.1; PMID:9995
 R:Smith, L.F.

Am. J. Med. 40, 662-666, 1966
 A:Title: Species variation in the amino acid sequence of insulin.
 A:Reference number: A90029; MUID:65160119; PMID:5949593
 A:Accession: A01587
 A:Molecule type: protein
 A:Residues: 25-54; 90-110 <SMI>
 F:Peterson, J.D.; Nehrlich, S.; Over, P.E.; Steiner, D.F.
 J. Biol. Chem. 247, 4866-4871, 1972
 A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsulins
 A:Reference number: A92111; MUID:72258016; PMID:4626369
 A:Accession: S16493
 A:Molecule type: protein
 A:Residues: 65-85; 1'-87 <PEP>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BCH>
 F:55-89/Domain: insulin chain A #status experimental <MAT>
 F:90-110/Domain: insulin connecting C peptide #status predicted <CPEP>
 F:91-96/Domain: insulin chain A #status experimental <ACH>
 F:91-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 90.1% Score 417; DB 1; Length 110;

Best Local Similarity 89.5% Pred. No. 2.9e-38;

Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 ALQKRGIVEQCCTISCSLYOLENYCN 110

RESULT 7

IPNG

insulin precursor - horse
 C:Species: Equus caballus (domestic horse)
 C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
 C:Accession: A01580; A92120
 R:Harris, J.J.; Sanger, F.; Naughton, M.A.
 Arch. Biochem. Biophys. 65, 427-428, 1956
 A:Title: Species differences in insulin.
 A:Reference number: A90082
 A:Accession: A01580
 A:Molecule type: protein
 A:Residues: 1-30; 66-86 <HAR>
 R:Tager, H.S.; Steiner, D.F.
 J. Biol. Chem. 247, 7916-7940, 1972
 A:Title: Primary structures of the proinsulin connecting peptides of the rat and human insulin
 A:Reference number: A92120; MUID:73061498; PMID:4640931

Query Match 90.1% Score 417; DB 1; Length 110;

Best Local Similarity 89.5% Pred. No. 2.9e-38;

Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVQLVGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 ALQKRGIVEQCCTISCSLYOLENYCN 110

```
A:Accession: A92120
A:Molecule type: protein
A:Residues: 33-63 <TAG>
C:Comment: X's at positions 31-33 and 64-65 represent paired basic residues assumed (by
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <BCH>
F:31-30/Domain: insulin chain B #status experimental <BCH>
F:33-63/Domain: connecting peptide #status experimental <CPEP>
F:66-86/Domain: insulin chain A #status experimental <ACH>
F:7-72,79-85,71-75/Disulfide bonds: #status predicted

Query Match      85.1%: Score 394; DB 1; Length 86;
Best Local Similarity 84.9%: Pred. No. 9.2e-36;
Matches 73; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKAXXAEDLQVGQVELGGGPGAGSQPLALEG 60
OY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 61 FQXXAGIVEQCCTSCSLYLENYCN 86

RESULT 8
Insulin 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 21-Sep-1999
C:Accession: B90789; B94231; C91231; I64863; A01530; B42120
R:Lomedico, P.; Rosenthal, A.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.;
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A:Reference number: A90789; MUID:8004535; PMID:458284
A:Accession: B90789
A:Molecule type: DNA
A:Residues: 1-110 <LOM>
A:Cross-references: G5:200748; NID:q204958; PIDN:AAA41443.1; PID:q204958
R:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Matquias, E.; Aron, H.; Oyler
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; MUID:7096763; PMID:4311936
A:Accession: B94231
A:Molecule type: protein
A:Residues: 25-54; 90-110 <STE>
R:Tagger, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A:Reference number: A92120; MUID:73061498; PMID:4645931
A:Accession: C92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Lomedico, P.T.; Rosenthal, A.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N.Y. Acad. Sci. 343, 425-432, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: I51945; MUID:80240379; PMID:6249167
A:Accession: I51945
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M25595; NID:q204950; PIDN:AAA41443.1; PID:q204952
C:Genetics:
A:Gene: INS2
A:Introns: 63/1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:55-90-110/Domain: insulin chain A #status experimental <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status experimental

A:Accession: A92120
A:Molecule type: protein
A:Residues: 33-63 <TAG>
C:Comment: X's at positions 31-33 and 64-65 represent paired basic residues assumed (by
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <BCH>
F:31-30/Domain: insulin chain B #status experimental <BCH>
F:33-63/Domain: connecting peptide #status experimental <CPEP>
F:66-86/Domain: insulin chain A #status experimental <ACH>
F:7-72,79-85,71-75/Disulfide bonds: #status predicted

Query Match      85.1%: Score 394; DB 1; Length 110;
Best Local Similarity 84.9%: Pred. No. 9.2e-36;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPMSRREVDPQVAQLFELGGGPGAGDLOTLEAV 84
OY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 61 AQKRGIVDQCCTSCSLYLENYCN 110

RESULT 9
Insulin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
C:Accession: A26342; B48172; A61012; B01592
R:Kotwirth, H.M.; Schaefer, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.
J. Mol. Evol. 23, 305-312, 1986
A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
A:Reference number: A92965; MUID:87169768; PMID:3104603
A:Accession: A26342
A:Molecule type: DNA
A:Residues: 1-110 <MEN>
A:Cross-references: GB:X04724; NID:q52714; PIDN:CAA28433.1; PID:q52715
R:Swia, T.; Ogata, S.; Morioka, H.; Yang, S.
J. Mol. Endocrinol. 5, 61-67, 1990
A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the
A:Reference number: A48172; MUID:90372989; PMID:2337023
A:Accession: B48172
A:Status: not compared with conceptual translation
A:Molecule type: DNA
R:Linde, S.; Nielsen, J.H.; Hansen, B.; Wellinder, B.S.
J. Chromatogr. 462, 243-254, 1989
A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin
A:Reference number: A61012; MUID:89292078; PMID:2661585
A:Accession: A61012
A:Molecule type: protein
A:Residues: 57-87 <LIN>
R:Buenzli, H.F.; Glatthaar, B.; Kienz, P.; Muelhaupt, E.; Humbel, R.E.
Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
A:Reference number: A01592; MUID:72189455; PMID:5063718
A:Accession: B01592
A:Molecule type: protein
A:Residues: 25-54; 90-110 <BOR>
C:Genetics:
A:Introns: 63/1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54,90-110/Domain: insulin chain A #status experimental <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match      85.1%: Score 394; DB 1; Length 110;
Best Local Similarity 84.9%: Pred. No. 9.2e-36;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPMSRREVDPQVAQLFELGGGPGAGDLOTLEAV 84
OY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 61 AQKRGIVDQCCTSCSLYLENYCN 110
```

```

RESULT 10
A:Species: 1-110 <CCR>
A:Cross-references: GR:J00747; NID:9204956; PIDN:AAA41442.1; PID:g204957
A:Medication: F.; Rosenzweig, A.; Efratiadis, A.; Gilbert, W.; Kolodner, R.; Lizard
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes
A:Reference number: A90789; MUID:80045034; PMID:498284
A:Accession: A39883
A:Molecule type: DNA
A:Residues: 1-110 <DOM>
A:Cross-references: GR:J00747; NID:9204956; PIDN:AAA41442.1; PID:g204957
A:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margolis, E.; Allen, B
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; MUID:70067613; PMID:4311936
A:Accession: A94231
A:Molecule type: Protein
A:Residues: 25-54:90-110 <STE>
R:Tagar, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and h
A:Reference number: A92120; MUID:73061498; PMID:464093
A:Accession: B92120
A:Molecule type: Protein
A:Residues: 57-87 <TAG>
R:Medico, P.; Rosenthal, N.; Kolodner, R.; Efratiadis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-432, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: 151945; MUID:80240379; PMID:6249167
A:Accession: 151945
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M25584; NID:9204947; PIDN:AAA41439.1; PID:g204948
C:Genetics:
A:Gene: INS1
C:Superfamily: Insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <RCH>
F:57-87/Domain: connecting peptide #status experimental <MAP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-95:43-109,95-100/Disulfide bonds: #status experimental
Query Match 84.7% Score 385; DB 1; Length 110;
Best Local Similarity 83.7% Pred. No. 8; Re-35;
Matches 73; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVENLVCCGEGFYTPKTRAEACLVGVVHLCGGGACGASLOPLALEG 60
DB 25 FVNHLCGSHLVENLVCCGEGFYTPKTRAEACLVGVVHLCGGGACGASLOPLALEG 60
QY 61 SLOKRGIVCCCTSCSYQLONYCN 86
DB 83 PMKRCVDDCCCTSCSYQLONYCN 108

RESULT 11
insulin precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16 Jul-1999
C:Accession: 148166
R:Bell, G.I.; Sanchez-Pescador, R.
Diabetes 33, 297-305, 1984
A:Title: Sequence of a cDNA encoding Syrian hamster preproinsulin.
A:Reference number: 148166; MUID:84133036; PMID:6365563
A:Accession: 148166
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M26328; NID:4194420; PIDN:AAA37089.1; PID:g307360
C:Superfamily: Insulin

Query Match 84.7% Score 372; DB 2; Length 110;
Best Local Similarity 84.9% Pred. No. 1; Re-35;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVENLVCCGEGFYTPKTRAEACLVGVVHLCGGGACGASLOPLALEG 60
DB 25 FVNHLCGSHLVENLVCCGEGFYTPKTRAEACLVGVVHLCGGGACGASLOPLALEG 60
QY 61 SLOKRGIVCCCTSCSYQLONYCN 86
DB 85 AOKRGIVCCCTSCSYQLONYCN 110

RESULT 12
insulin 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 24 Sep-1999
C:Accession: A90788; A90789; A94231; B92120; 151945; A01589
R:Cordell, B.; Bell, G.; Tischer, E.; DeNoto, F.M.; Ellis, R.M.; Broder, W.W.
Cell 18, 533-543, 1979
A:Title: Isolation and characterization of a cloned rat insulin gene.
A:Reference number: A90788; MUID:80045034; PMID:498283
A:Accession: A90788
A:Molecule type: DNA

```



```

A:Accession: A94572
A:Molecule type: protein
A:Residues: 1-84 <CH2>
R:Brown, H.; Sanger, F.; Kitai, R.
Biochem. J. 60, 556-565, 1955
A:Title: The structure of pig and sheep insulins.
A:Reference number: A90344
A:Accession: S16492
A:Molecule type: protein
A:Residues: 1-30:31-51 <BR>
R:Snell, L.; Damgaard, U.
Horm. Metab. Res. 20, 476-480, 1988
A:Title: Proinsulin heterogeneity in pigs.
A:Reference number: A60635; MUID:89032178; PMID:3141865
A:Accession: A60535
A:Molecule type: protein
A:Residues: 33-62 <SN2>
A:Note: the authors report the characterization of a connecting peptide variant lacking
A:Accession: B62835
A:Molecule type: protein
A:Residues: 33-62 <SN2>
R:Blundell, T.; Dodson, G.; Hodgkin, D.; Moras, D.
Adv. Protein Chem. 26, 279-402, 1972
A:Title: Insulin: the structure in the crystal and its reflection in chemistry and biology
A:Reference number: A90017
A:Contents: annotation; X-ray crystallography, 1.9 Angstroms
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <EXP>
F:1-30:64-84/Product: insulin #status experimental <EXP>
F:33-63/Domain: connecting peptide #status experimental <CPEP>
F:64-84/Domain: insulin chain A #status experimental <ACH>
F:7-70:19-83:69-74/Disulfide bonds: #status experimental

Query Match 82.7% Score 383; DB 1; Length 84;
Best Local Similarity 86.0%; Pred. No. 1,16-14;
Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

QY : FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGPAGSLOQLALEG 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGPAGSLOQLALEG 84

QY 61 SLQKRGIVGQCCTSCISLYOLENYCN 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 59 PPQKRGIVGQCCTSCISLYOLENYCN 84

RESULT 14
IPBO
Insulin precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence_revision 22-Apr-1995 #text_change 16-Jul-1999
C:Accession: A40909; A92080; A92074; A91185; A90342; A90341; S48184; S46248; A01
R:D'Agostino, J.; Younes, M.A.; White, J.W.; Bosch, P.K.; Field, J.S.; Frazier, M.L.
Mol. Endocrinol. 1, 327-331, 1987
A:Title: Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid
A:Reference number: A40909; MUID:88288209; PMID:2456452
A:Accession: A40909
A:Molecule type: mRNA
A:Residues: 1-105 <DAA>
A:Cross-references: GB:M54573; NID:g163573; PDB:AAA30722.1; PDB:q163573
A:Experimental source: fetal pancreas
R:Nolan, C.; Margoliash, E.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 2780-2795, 1971
A:Title: The structure of bovine proinsulin.
A:Reference number: A92080; MUID:71166442; PMID:4928932
A:Accession: A92080
A:Molecule type: protein
A:Residues: 25-105 <NCL>
R:Steiner, D.F.; Cho, S.; Oyer, P.E.; Terris, S.; Peterson, J.D.; Krausstein, A.H.
J. Biol. Chem. 245, 1365-1374, 1971
A:Title: Isolation and characterization of proinsulin C-peptide from bovine pancreas.
A:Reference number: A92074; MUID:71166409; PMID:5545080

```

```

A:Accession: A92074
A:Molecule type: protein
A:Residues: 57-82 <STG>
R:Salokangas, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 183-189, 1971
A:Title: Bovine proinsulin: amino acid sequence of the C-peptide isolated from panc
A:Reference number: A91185; MUID:71257721; PMID:5105368
A:Accession: A91185
A:Molecule type: protein
A:Residues: 57-82 <SAI>
R:Sanger, F.; Thompson, E.G.P.
Biochem. J. 53, 366-374, 1953
A:Title: The amino-acid sequence in the glycol chain of insulin. 2. The investigati
A:Reference number: A90342
A:Accession: A90342
A:Molecule type: protein
A:Residues: 85-105 <SAN>
R:Sanger, F.; Tuppy, H.
Biochem. J. 49, 481-490, 1951
A:Title: The amino-acid sequence in the phenylalanyl chain of insulin. 2. The inves
A:Reference number: A90341
A:Accession: A90341
A:Molecule type: protein
A:Residues: 25-54 <SA2>
R:Cheng, R.; Kawakishi, S.
Eur. J. Biochem. 223, 759-764, 1994
A:Title: Site-specific oxidation of histidine residues in glycosylated insulin mediated
A:Reference number: S48184; MUID:94333378; PMID:8055951
A:Accession: S48184
A:Molecule type: protein
A:Residues: 85-105 <CHE>
A:Accession: S48195
A:Status: preliminary
A:Molecule type: protein
A:Residues: 25-30; X' 32-42; X' 44-54 <CH2>
R:Kryle, A.P.; Sanger, F.; Smith, L.F.; Kitai, R.
Biochem. J. 60, 541-556, 1955
A:Title: The disulphide bonds of insulin.
A:Reference number: A90343
A:Contents: annotation; acids; disulfides
R:Kendall, T.; Eckerskorn, C.; Lottspeich, F.; Baummeister, W.
FEBS Lett. 349, 205-209, 1994
A:Title: Existence of a molecular ruler in proteasomes suggested by analysis of deg
A:Reference number: S46258; MUID:94326921; PMID:8050567
A:Accession: S46258
A:Status: preliminary
A:Molecule type: protein
A:Residues: 25-54 <WEN>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54:95-105/Product: insulin #status experimental <WCH>
F:57-82/Domain: connecting peptide #status experimental <CPEP>
F:85-105/Domain: insulin chain A #status experimental <ACH>
F:31-91:43-104:90-95/Disulfide bonds: #status experimental

Query Match 79.2% Score 366.5; DB 1; Length 105;
Best Local Similarity 80.2%; Pred. No. 8,7e-33;
Matches 69; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

QY 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGPAGSLOQLALEG 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGPAGSLOQLALEG 79

QY 61 SLQKRGIVGQCCTSCISLYOLENYCN 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 PPQKRGIVGQCCTSCISLYOLENYCN 105

RESULT 15
IPMSI
Insulin 1 precursor - mouse

```

C:Species: Mus musculus (house mouse)
 C:Date: 24-Apr-1984 #sequence_revision 11-Jul-1994 #text_change 12-Jul-1994
 C:Accession: 926342; A01592; B61012
 R:Wentworth, B.M.; Schaefer, T.M.; Villa-Komaroff, L.; Chiriac, J.R.
 J. Mol. Evol. 23, 305-312, 1986
 A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
 A:Reference number: A92965; MUID:8716978; PMID:3704833
 A:Accession: 926342
 A:Molecule type: DNA
 A:Residues: 1-108 <GEN>
 A:Cross-references: GH:G04725; NID:952712; PIDN:CAA28434.1; PID:952713
 R:Sawa, T.; Otsuka, S.; Morioka, H.; Yano, S.
 J. Mol. Endocrinol. 5, 61-67, 1990
 A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the R.R.M.
 A:Reference number: A48172; MUID:90372989; PMID:2397023
 A:Accession: A48172
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-108 <SAW>
 R:Buenzli, H.F.; Glatthard, R.; Kitz, P.; Muelhaupt, E.; Humbel, R.E.
 Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
 A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
 A:Reference number: A01592; MUID:72189455; PMID:5063713
 A:Accession: A01592
 A:Molecule type: protein
 A:Residues: 25-54:88-108 <BOE>
 R:Linde, S.; Nielsen, J.H.; Hansen, B.; Welinder, B.S.
 J. Chromatogr. 462, 243-254, 1989.
 A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin bind
 A:Reference number: A61012; MUID:89292078; PMID:2651585
 A:Accession: B61012
 A:Molecule type: protein
 A:Residues: 57-85 <Lin>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BCB>
 F:25-54,88-108/Product: insulin #status experimental <MAI>
 F:57-85/Domain: connecting peptid #status experimental <CEP>
 F:88-108/Domain: insulin chain A #status experimental <ACH>
 F:31-94,43-107,93-98/Disulfide bonds: #status predicted

Query Match 79.9% Score 366; DB 1; Length 108;
 Best Local Similarity 81.4% Prod. No. 1212;
 Matches 70; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

OY 1 FVNCILGSGHVEALVYPCGFGKFFTPKTRPRRRLGQVGVGVGGGPRRANSGFALSG 60
 Db 25 FVQHEGSPHVEALVYCGRGFFTPKSRREVEVQVELEGGSP--GGLQTALREV 82
 OY 61 SLAKRGIVEGCTSIGS YDLENVCN 86
 Db 83 ARQKRGIVDCCISGNYQLENVCN 108

Search completed: September 15, 2003, 12:04:19
 Job time : 14.3792 secs

Genome version 5.1.6
Copyright (c) 1993 - 2003 Computer Sci.

OM protein - protein search, us sw model

Run on: September 15, 2003, 11:54:00 : Search time 4:01:34 Seconds
(without alignments)
564,633 Million cell updates/sec

Title: US-09-423-100-4

Perfect score: 463

Sequence: 1 FVQVHLCGSHVEALRWG.....VFGQTSICSLVQLRHHN 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47626705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 260000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	463	100.0	110	INS_HUMAN	201309 Homo sapien
2	463	100.0	110	INS_PANTR	P03110 Pan troglod
3	456	98.5	110	INS_GORAE	P04746 Gorilla gorilla
4	456	98.5	110	INS_MACFA	P04746 Macaca fasci
5	424	91.6	110	INS_RABIT	P01311 Caprolagus
6	417	90.1	110	INS_CANFA	P01321 Canis fam
7	413	89.2	110	INS_SPERF	Q31313 Spermophila
8	394	85.1	86	INS_BORSE	P01310 Sus scrofa
9	394	85.1	110	INS2_MOUSE	P01326 Mus musculus
10	392	84.7	110	INS2_RAT	P01325 Rattus norv
11	392	84.7	108	INS2_OTTK	P12604 Octopus tigr
12	392	84.7	110	INS_CRILLO	P01313 Crinoidus
13	385	83.2	110	INS1_RAT	P01322 Rattus norv
14	383	82.7	108	INS_PIG	P01315 Sus scrofa
15	377	81.4	110	INS_PSAPE	Q62567 Psammis o
16	366.5	79.2	105	INS_BOVIN	P01317 Bos taurus
17	366	79.0	108	INS1_MOUSE	P01321 Mus musculus
18	362.5	78.3	105	INS_SHEEP	P01318 Ovis aries
19	334.5	72.2	108	INS2_MOUSE	P21553 Rattus norv
20	314	67.4	110	INS_CAVIO	P01323 Canis lupus
21	277.5	59.5	105	INS2_OTTK	P01312 Octopus oeg
22	276.5	59.7	103	INS2_BELF	P01314 Sepioides
23	265.5	57.3	105	INS2_XENLA	P12732 Xenopus lae
24	265.5	57.3	107	INS2_CHICK	P01322 Gallus gall
25	262.5	56.7	106	INS1_XENLA	P12732 Xenopus lae
26	256.5	55.4	51	INS2_BALPH	P01312 Balanophor
27	256.5	55.4	51	INS1_ELEVA	P01316 Elephas max
28	256.5	55.4	81	INS1_WAP2	P01333 Anas platyr
29	251.5	54.3	51	INS1_ACOCA	P01324 Accipiter
30	246.5	53.2	51	INS1_BALBO	P01324 Balaeophor
31	246.5	53.2	51	INS1_CAMDR	P01320 Camelus dfo
32	246.5	53.2	51	INS1_CAPRI	P01319 Capra hirc
33	245.5	53.0	51	INS1_FELCA	P06306 Felis silve

ALIGNMENTS				
RESULT 1				
INS_HUMAN	STANDARD	PRT:	110 AA.	
AC P01309;				P01309 didelphis m
DT 21-JUL-1986 (Rel. 02, Created)				P01327 chinchilla
DT 21-JUL-1986 (Rel. 01, Last sequence update)				P07454 anser anser
DT 15-SEP-2003 (Rel. 42, Last annotation update)				Q9W752 verasper mo
CE Insulin precursor.				P01328 hystrix cri
GN INS.				P31857 trachemys s
OS Homo sapiens (Human).				P01423 acipenser g
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				P01341 lophius pis
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				Q9W757 ornithorhyn
OX NCBI_TaxID=9606;				P04667 oncorhynch
RP SEQUENCE FROM N.A.				P01306 callorhynch
RX MEDLINE=80120725; PubMed=6243748;				P01025 oreochromis
RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischler E.,				
RA Goodman H.M.;				
RT "Sequence of the human insulin gene.";				
RL Nature 284:26-32(1985).				
RN [1]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=80236133; PubMed=6248962;				
RA Ulrich A., Dull T.J., Gray A., Brosius J., Sures I.,				
RA "Genetic variation in the human insulin gene.";				
RL Science 209:612-615(1983).				
RN [3]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=80054779; PubMed=503234;				
RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,				
RA Rutter W.J.;				
RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin.";				
RL Nature 262:525-527(1979).				
RN [4]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=80147417; PubMed=6927840;				
RA Sures I., Goeddel D.V., Gray A., Ulrich A.;				
RA "Nucleotide sequence of human preproinsulin complementary DNA.";				
RL Science 208:57-59(1980).				
RN [5]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=93364428; PubMed=8358440;				
RA Lussen A.M., Bell T.J., Julier C., Latour M.;				
RA "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1				
PT kb segment of DNA spanning the insulin gene and associated VNTR.";				
RL Nat. Genet. 4:305-310(1993).				
RN [6]				
RP SEQUENCE FROM N.A.				
OC TISSUE=pacreas;				
RX MEDLINE=22388257; PubMed=12477932;				
RA Strausberg R., Feingold E.A., Grouse L.H., Berge J.G.,				
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA Altschul S.F., Zeeberg K., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,				
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				

RA Brownstein M.J., Ustin T.B., Teshiyuki S., Carninci P., Prange G.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mellichy S.C.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Guinetone P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay J., Hulyk S.W.,
RA Villalon D.K., Muzay D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Ketteran M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard J.G.,
RA Blakesley R.W., Touchman J.W., Grein E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,
RA Butterfield V.S.N., Krzywinski M.J., Skalska H., Smallos D.E.,
RA Schnerch A., Schell J.E., Jones S.J.M., Karra R.A.,
RT "Generation and initial analysis of more than 55,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
RN [7]
RN TISSUE=Blood;
RC SEQUENCE OF 1-59 FROM N.A.
RA Fajardy I., Weill J.J., Stuckens C.C., Danze P.M.F.;
RT "Description of a novel RFLP diallelic polymorphism (-127 HpaI G/G)
RT within the 5' region of insulin gene.";
RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ Databases.
RN [8]
RN SEQUENCE OF 25-54 AND 90-110.
RA Nicol D.S.H., Smith I.F.;
RT "Amino-acid sequence of human insulin.";
RJ Nature 187:483-485(1966).
RN [9]
RN SEQUENCE OF 57-87.
RX MEDLINE=71257722; PubMed=5563484;
RA KC A., Smyth D.G., Markusen J., Sindby F.;
RT "The amino acid sequence of the C-peptide of human proinsulin.";
RJ Eur. J. Biochem. 20:199-199(1971).
RN [11]
RN SYNTHESIS.
RX MEDLINE=75077277; PubMed 444429;
RA Steber P., Kamber B., Hartmann A., Jeoll A., Rinke B., Ritzel W.,
RT "Total synthesis of human insulin under directed formation of the
RT disulfide bonds.";
RJ Helv. Chim. Acta 57:2617-2621(1974).
RN [12]
RN SYNTHESIS OF 57-87.
RX MEDLINE=75040007; PubMed=4833504;
RA Naithani V.K.;
RT "Studies on polypeptides, IV. The synthesis of C-peptide of human
RT proinsulin.";
RJ Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
RN [13]
RN SYNTHESIS OF 65-69 AND 70-73.
RX MEDLINE=73161263; PubMed=4698555;
RA Geiger R., Volk A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 6-13
RT of human proinsulin C peptides.";
RJ Chem. Ber. 106:199-205(1973).
RN [14]
RN SYNTHESIS OF 84-87.
RX MEDLINE=73161261; PubMed=4698553;
RA Geiger R., Geiger G., Keck W., Treuth G.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 1. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide.";
RJ Chem. Ber. 106:186-192(1973).
RN [15]
RN VARIANT: LOS ANGELES SER-48.
RX MEDLINE=84016053; PubMed=6312459;
RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
RT specific resonance assignments and effects of solvent composition.";
RJ Biochim. Biophys. Acta 1078:101-110(1991).
RN [16]
RN "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB24]insulin.";
RJ Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
RN [18]
RN VARIANTS LOS ANGELES SER-46 AND CHICAGO LEU-49.
RX MEDLINE=84170233; PubMed=642411;
RA Shewelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.;
RT "Identification of a mutant human insulin: predicted to contain a
RT serine-for-phenylalanine substitution.";
RJ Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RN [17]
RN VARIANT PROVIDENCE ASP-34.
RX MEDLINE=87175640; PubMed=3470784;
RA Chan S.J., Seino S., Gruppuso P.A., Schwartz R., Steiner D.F.;
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RJ Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
RN [18]
RN VARIANT WAKAYAMA LEU-92.
RX MEDLINE=87058122; PubMed=3537011;
RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A3 (Val->Leu) isolated from the pancreas.";
RJ J. Clin. Invest. 78:1668-1672(1986).
RN [19]
RN VARIANT HIS-89.
RX MEDLINE=90317021; PubMed=2196379;
RA Barbetti F., Raben N., Katswaki T., Cama A., Accili D., Gabbay K.H.,
RA Meyerich J.A., Taylor S.J., Roth Y.;
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RJ J. Clin. Endocrinol. Metab. 71:164-169(1990).
RN [20]
RN VARIANT HIS-89.
RX MEDLINE=85261996; PubMed=4019786;
RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RJ J. Clin. Invest. 76:374-380(1985).
RN [21]
RN VARIANT KYOTO LEU-89.
RX MEDLINE=92291307; PubMed=1601597;
RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Izura H., Seino Y.;
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto).";
RJ J. Clin. Invest. 89:1902-1907(1992).
RN [22]
RN STRUCTURE BY NMR.
RX MEDLINE=91104956; PubMed=2271664;
RA Hua Q.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RT structure.";
RJ Biochemistry 29:10545-10555(1990).
RN [23]
RN STRUCTURE BY NMR.
RX MEDLINE=91242467; PubMed=2036420;
RA Hua Q.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition.";
RJ Biochemistry 30:5505-5515(1991).
RN [24]
RN STRUCTURE BY NMR.
RX MEDLINE=91265527; PubMed=1546635;
RA Hua Q.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
RT specific resonance assignments and effects of solvent composition.";
RJ Biochim. Biophys. Acta 1078:101-110(1991).

```
Query Match      100.0%; Score 463; 25 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 9.2e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Query Match      98.5%   Score 456;   DR 1;   Length 110;
Best Local Similarity 98.8%   Pred. No. 5,2e-42;
Matches 85;   Conservative 0;   Mismatches 1;   Indels 6;   Gaps 0;

QY 1 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGGLGGPGAGSLQPLALEG 60
DB 25 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGGLGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVECCCTSCISLYQLENYCN 86
DB 85 SLQKRGIVECCCTSCISLYQLENYCN 110

RESULT 5
INS_PABIT STANDARD: PRT: 110 AA.
AC P01311;
DT 21-JUL-1986 (Rel. 01; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 01-OCT-1996 (Rel. 33; Last annotation update)
DE Insulin precursor.
GN INS.
OS Crystallagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Crystallagus.
CX NCBI_TaxID=9956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8308074; PubMed=6164262;
RA Winkler E.L.;
RT "The nucleotide sequence of cDNA coding for preproinsulin from the
RL gene 13:19-183(1982).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J00336; AAA16849.1;
CC PIR: J00178; J00178.
CC HSP: P01308; IAIC.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00649; Insulin;
CC PRINTS: PR00277; INSULINH.
CC SMART: SM00078; IIGF;
CC PROSITE: PS00262; INSULIN;
CC Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
SQ SEQUENCE 110 AA; 8306ET3A63A420E9 CDS64;

Query Match      98.5%   Score 456;   DR 1;   Length 110;
Best Local Similarity 98.8%   Pred. No. 5,2e-42;
Matches 85;   Conservative 0;   Mismatches 1;   Indels 6;   Gaps 0;

QY 1 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGGLGGPGAGSLQPLALEG 60
DB 25 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGGLGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVECCCTSCISLYQLENYCN 86
DB 85 SLQKRGIVECCCTSCISLYQLENYCN 110

RESULT 5
INS_PABIT STANDARD: PRT: 110 AA.
AC P01311;
DT 21-JUL-1986 (Rel. 01; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 01-OCT-1996 (Rel. 33; Last annotation update)
DE Insulin precursor.
GN INS.
OS Crystallagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Crystallagus.
CX NCBI_TaxID=9956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8308074; PubMed=6164262;
RA Winkler E.L.;
RT "The nucleotide sequence of cDNA coding for preproinsulin from the
RL gene 13:19-183(1982).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J00336; AAA16849.1;
CC PIR: J00178; J00178.
CC HSP: P01308; IAIC.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00649; Insulin;
CC PRINTS: PR00277; INSULINH.
CC SMART: SM00078; IIGF;
CC PROSITE: PS00262; INSULIN;
CC Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
SQ SEQUENCE 110 AA; 8306ET3A63A420E9 CDS64;

```

FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 E - Y (IN REF. 8)
SQ SEQUENCE 110 AA: 11838 MW: 82029/53855/7788 CRC64:

Query Match 91.54; Score 424; DB 1; Length 110;
Best Local Similarity 90.74; Pred. No. 140-38.
Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNHGCSHLVLAALYLVCGERGFYTPKTRKREAEALQVQVLEGGGAGSLQPLALEG 60
DQ 25 FVNHGCSHLVLAALYLVCGERGFYTPKTRKREAEALQVQVLEGGGAGSLQPLALEG 64

QY 61 SLQKRGIVFQCCSTCSLYQLENYCN 86
DQ 85 ALQKRGIVFQCCSTCSLYQLENYCN 110

RESULT 6
INS_CANFA STANDARD; PRT: 110 AA.
AC P01321;
DT 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-FEB-1996 (Rel. 03, Last annotation update)
GN INS
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83109071; PubMed=6296142;
RA Kwok S.C.M., Chan S.J., Steiner D.P.;
RT "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT Coded amino acid sequence of canine preproinsulin products an
RT additional C-peptide fragment."
RL J. Biol. Chem. 258:2357-2364(1983).
RN [2]
RP SEQUENCE OF 25 54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949503;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V00179; CAA23475.1;
CC PIR: A92413; IPDG.
CC HSSP: P01317; IAPB.
CC InterPro: IPR004925; Ins/IGF/relax.
CC Pfam: PF00649; Insulin; 1.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.
FT DISULFID 31 96
FT DISULFID 43 109
FT DISULFID 95 100
FT CONFLICT 83 E - Y (IN REF. 8)
SQ SEQUENCE 110 AA: 11838 MW: 82029/53855/7788 CRC64:

FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12190 MW: 45747918644/4F98 CRC64:

Query Match 90.14; Score 417; DB 1; Length 110;
Best Local Similarity 89.58; Pred. No. 770-38.
Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNHGCSHLVLAALYLVCGERGFYTPKTRKREAEALQVQVLEGGGAGSLQPLALEG 60
DQ 25 FVNHGCSHLVLAALYLVCGERGFYTPKTRKREAEALQVQVLEGGGAGSLQPLALEG 84

QY 61 SLQKRGIVFQCCSTCSLYQLENYCN 85
DQ 85 ALQKRGIVFQCCSTCSLYQLENYCN 110

RESULT 7
INS_SPETR STANDARD; PRT: 110 AA.
AC Q91X13;
DI 15-SEP-2003 (Rel. 42, Created)
DI 15-SEP-2003 (Rel. 42, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Spromophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spromophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83109071; PubMed=6296142;
RA Kwok S.C.M., Chan S.J., Steiner D.P.;
RT "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT Coded amino acid sequence of canine preproinsulin products an
RT additional C-peptide fragment."
RL J. Biol. Chem. 258:2357-2364(1983).
RN [2]
RP SEQUENCE OF 25 54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949503;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AY038604; AAK72558.1;
CC HSSP: P01308; IAPB.
CC InterPro: IPR004925; Ins/IGF/relax.
CC Pfam: PF00649; Insulin; 1.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.
FT DISULFID 31 96
FT DISULFID 43 109
FT DISULFID 95 100
FT CONFLICT 83 E - Y (IN REF. 8)
SQ SEQUENCE 110 AA: 12004 MW: 45117680662/2BEE5 CRC64:

Cell 18:593-542(1979).
[2]
SEQUENCE FROM N.A.
RC STRAIN-SQUAMC-Dawley; T18-DB-11901;
RA MEDLINE=60949035; PubMed=45424;
RA Lomedico P., Rosenthal N., Estradiadis A., Gilbert W., Klotzner P.,
RA Tizard R.;
KT "The structure and evolution of the two homologous rat preproinsulin
genes";
RL Cell 18:545-558(1979).
[3]
SEQUENCE FROM N.A.
RP MEDLINE=80240379; PubMed=6249157;
RA Lomedico P.L., Rosenthal N., Kozloski R., Estradiadis A.,
RA Gilbert W.;
RT "The structure of rat preproinsulin genes";
RL Ann. N.Y. Acad. Sci. 343:425-432(1980).
[4]
SEQUENCE OF 25-54 AND 90-110.
RP MEDLINE=70067613; PubMed=4111935;
RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Matzoulis E.,
RA Aten B., Oyer P.E.;
RT "Proinsulin and the biosynthesis of insulin";
RL Recent Prog. Horm. Res. 25:207-282(1969).
[5]
SEQUENCE OF 57-87.
RP MEDLINE=73061499; PubMed=4640331;
RA Tager H.S., Steiner D.F.;
RT "Primary structures of the proinsulin connecting peptides of the rat
and the horse";
RL J. Biol. Chem. 247:7935-7940(1972).
[6]
SEQUENCE OF 57-87. AND REV-STONS.
RP MEDLINE=72177395; PubMed=4554104;
RA Markussen J., Sundby P.;
RT "Rat-proinsulin C-peptides. Amino-acid sequences";
RL Eur. J. Biochem. 25:153-162(1972).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A H CHAIN AND AN A CHAIN LINKED BY TWO
CC C-SULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE "INSULIN/IGF/RELAXIN FAMILY".
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.sib.ch/annouce/>
CC or send an email to license@isb.sib.ch).
CC
DR EMBL: V01242; CAA24559.1;
DR EMBL: J00747; AAA41442.1;
DR EMBL: M25584; AAA41439.1;
DR PIR: A30786; IPRT1.
DR HSSP: PC1308; 1A7F.
DR InterPro: IPR003425; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMO078; IIGF; 1.
DR PROSITE: PS02624; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN 1 B CHAIN.
FT PROPEP 57 87 INSULIN 1 C PEPTIDE.
FT CHAIN 90 110 INSULIN 1 A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12420 MW: 51D606DA54AE4533 CRC64:

Query Match 63.28; Score 385; DB 1; Length 110;
Best Local Similarity 83.7%; Pred. No. 2e-34;
Matches 72; Conservative 4; Mismatches 10; Gaps 0;
QY 1 FYNHLCGSHAEVLYVCGRGFYTPKTRAEEDLVGQVELGGGPGASSLQPLALEG 60
DE 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 25 FVQHQHCGPHEVLYVCGRGFFTPKSRREVEDQVPOLELGGPGAGPLQTLALIV 84
DE 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 61 SLOKSGIVEQCCTSCSLYQLENYCN 86
DE 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 85 ASQKRGIVEQCCTSCSLYQLENYCN 110
DE 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
RESULT 14
INS_P56
QY INS_P13 STANDARD; PRT: 108 AA.
DE 10-1315; 091535;
DE 21-JUL-1986 (Rel. 21. Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin precursor.
DE INS.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Han X.G., Tsch B.E.;
RT "Complete porcine proinsulin cDNA sequence";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Large white;
RX MEDLINE=22135958; PubMed=12140686;
RA Amarger V., Nguyen M., Laere A.S., Braunschweig M., Nezer C.,
RA Georges M., Andersson L.;
RT "Comparative sequence analysis of the INS-IGF2-H19 gene cluster in
RT pigs";
RL Mamm. Genome 13:388-398(2002).
RN [3]
RP SEQUENCE OF 25-108.
RX MEDLINE=6826485; PubMed=5657063;
RA Chance R.E., Ellis R.M., Bromer W.W.;
RT "Porcine proinsulin: characterization and amino acid sequence";
RL Science 161:165-167(1968).
RN [4]
RP REVISION TO 59.
RA Chance R.E.;
RL Submitted (JUL-1970) to the PIR data bank.
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RA Blundell T.L., Dodson G.G., Hodgkin D., Mercola D.;
RT "Insulin: the structure in the crystal and its reflection in
RT chemistry and biology";
RL Adv. Protein Chem. 26:279-402(1972).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA Isaacs N.W., Agarwal R.C.;
RT "Experience with fast Fourier least squares in the refinement of the
RT crystal structure of rhombohedral 2-zinc insulin at 1.5-A
RT resolution";
RL Acta Crystallogr. A 34:782-791(1978).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=89099316; PubMed=2905485;
RA Baker E.N., Blundell T.L., Cutfield S.M., Dodson E.J.,
RA Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
RA Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
RT "The structure of 22n pig insulin crystals at 1.5-A resolution";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-456(1988).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RX MEDLINE=92126280; PubMed=172613;
 RA Baischmidt P., Hansen P.B., Dodson E., Dodson G., Korber P.;
 RT "Structure of porcine insulin co-crystallized with chloride Z.";
 RL Acta Crystallogr. B 47:975-986(1991).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=9422450; PubMed=2025410;
 RA Badger J., Harris M.R., Reynolds C.D., Evans A., Dodson E.;
 RT "Structure of the pig insulin dimer in the cubic crystal.";
 RL Acta Crystallogr. B 47:127-136(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
 RA Diaio J.-S., Wan Z.-L., Chang W.-R., Liang D.-C.;
 RT "Structure of monomeric porcine desB1-B2 despancpeptide (E26-E30)
 in insulin at 1.65-A resolution.";
 RL Acta Crystallogr. D 53:507-512(1997).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC -!- DATABSE: NAME=Protein Spotlight;
 CC NOTE=Issue 9 of April 2003;
 CC WWW="http://www.expasy.org/spotlight/articles/spilt_09.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AF054555; AAC77920.1; ALI_NIT.
 DR EMBL: AY044828; AAL69555.1; -.
 DR PDB: 3INS; 09-JUN-99.
 DR PDB: 4INS; 31-JUL-94.
 DR PDB: 6INS; 31-JAN-94.
 DR PDB: 7INS; 31-JAN-94.
 DR PDB: 9INS; 15-OCT-91.
 DR PDB: 11ZA; 15-OCT-91.
 DR PDB: 11ZB; 15-OCT-91.
 DR PDB: 2IC1; 29-JAN-96.
 DR PDB: 1ME3; 29-JAN-96.
 DR PDB: 3KTH; 29-JAN-96.
 DR PDB: 1DE1; 16-JUN-97.
 DR PDB: 1SDB; 01-APR-98.
 DR PDB: 1WAF; 28-FEB-97.
 DR PDB: 12F1; 16-FEB-99.
 DR PDB: 12N1; 28-JAN-98.
 DR PDB: 12NJ; 26-JAN-98.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 95 C PEPTIDE.
 FT CHAIN 88 108 INSULIN A CHAIN.
 FT DISULFID 31 94 INTERCHAIN.
 FT DISULFID 43 107 INTERCHAIN.
 FT DISULFID 93 98
 FT HELIX 26 45
 FT STRAND 48 94
 FT HELIX 89 94
 FT HELIX 103 106
 FT STRAND 107 107
 FT SEQUENCE 108 AA; 1107 MF; CB4491H429858DE CRC64;

Query Match 82.7%; Score 383; DB 1; Length 108;
 Best Local Similarity 86.0%; Pred. No. 3.3e-34;
 Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;
 QY 1 FVNCHLGGSHLYVLCGERGFYTPKRYREAEQLQVQVHGGPGAGSLQPIALHG 60
 LQ 25 FVNCHLGGSHLYVLCGERGFYTPKRYREAEQVQVHGGPGAGSLQPIALHG 82
 QY 61 SLQKRGIVEOCCTSLCYOLENYCN 86
 LQ 83 PPKRGIVEOCCTSLCYOLENYCN 108
 RESULT 15
 INS_PSAOB STANDARD; PRT; 110 AA.
 ID INS_PSAOB
 AC Q625R7;
 DT 31-NOV-1997 (Rel. 25, Created)
 DT 31-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 35, Last annotation update)
 DE Insulin precursor.
 GN INS
 OS Psammomys obesus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Psammomys.
 CX NCBI_TaxID=48139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=97305250; PubMed=9166665;
 RA Kaiser N., Baillyes E.M., Schneider B.S., Cerasi E., Steiner D.F.,
 RA Burton J.C., Gross D.C.;
 RT "Characterization of the unusual insulin of Psammomys obesus, a
 RT rodent with nutrition-induced NIDDM-like syndrome.";
 RL Diabetes 46:953-957(1997).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X98241; CAA66897.1; -.
 DR HSP; P01308; IAI0.
 DR InterPro: IPR004825; Ins/IGF/re-ax.
 DR Pfam: PF00049; Insulin; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 31 96 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 95 100
 FT SEQUENCE 110 AA; 12324 MW; A006738E20579CB0 CRC64;
 Query Match 81.4%; Score 377; DB 1; Length 110;
 Best Local Similarity 81.4%; Pred. No. 1.5e-33;

Matches	70	Conservative	5	Mismatches	11	Indels	0	Gaps	0
QY	1	FVNHLCSSHLV	ALV	SGRGFF	YTPKTR	SAE	EL	QVQV	QVQV
Db	25	FVNHLCSSHLV	ALV	SGRGFF	YTPKTR	SAE	EL	QVQV	QVQV
QY	61	SLQKRGIV	CCCT	IS	SLVQ	ENYCN	86		
Db	85	ARQKRGIV	CCCT	IS	SLVQ	ENYCN	119		

Search completed: September 15, 2003, 12:01:30
 Job time : 9.01434 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

CM protein - protein search, using sw model

Run on: September 16, 2003, 11:04:30 Search time 13,5986 seconds
(with 1000 alignments)
606,522 Million cell operations

Title: US-09-423-100-4
Perfect score: 453
Sequence: 1 FVACHLGSHLVPAALVCG.....IVQGVCTSSLSYLENYCN 86

Scoring table: RLSUN62
Gapop 10.0, Capext 0.5

Searched: 830525 seqs, 254052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first: 45 summaries

- Database: SPTREMBL_23.1
- 1: sp.archaea:
 - 2: sp.bacteria:
 - 3: sp.fungi:
 - 4: sp.human:
 - 5: sp.invertebrate:
 - 6: sp.mammal:
 - 7: sp.mhc:
 - 8: sp.organelle:
 - 9: sp.phage:
 - 10: sp.plant:
 - 11: sp.podent:
 - 12: sp.virus:
 - 13: sp.verticillate:
 - 14: sp.uncclassified:
 - 15: sp.virus:
 - 16: sp.bacteriophage:
 - 17: sp.archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	EB	ID	Description
1	463	100.0	110	6	Q8HXV2	Q8HXV2 protein, Homo
2	413	89.2	110	11	Q91X13	Q91X13 protein, Homo
3	388	83.8	110	6	Q8WAK6	Q8WAK6 protein, Homo
4	342	73.9	65	6	Q8WAK6	Q8WAK6 protein, Homo
5	342	73.9	65	6	Q8WAK6	Q8WAK6 protein, Homo
6	245.5	53.2	106	13	Q91807	Q91807 protein, Homo
7	235.5	50.9	111	13	Q98TA7	Q98TA7 protein, Homo
8	230.5	49.8	110	13	Q98TA7	Q98TA7 protein, Homo
9	222.5	48.1	120	13	Q98TA7	Q98TA7 protein, Homo
10	219	47.3	111	13	Q98TA7	Q98TA7 protein, Homo
11	214.5	46.3	108	13	Q98TA7	Q98TA7 protein, Homo
12	212.5	45.9	108	13	Q98TA7	Q98TA7 protein, Homo
13	210.5	45.5	87	13	Q98TA7	Q98TA7 protein, Homo
14	205.5	44.4	108	13	Q98TA7	Q98TA7 protein, Homo
15	203.5	44.0	91	13	Q98TA7	Q98TA7 protein, Homo
16	189	40.8	41	11	Q62543	Q62543 protein, Homo

17	162	35.0	75	11	Q62542	Q62542 protein, Homo
18	142.5	30.8	159	13	Q93607	Q93607 protein, Homo
19	142.5	30.8	142	13	Q93607	Q93607 protein, Homo
20	142.5	30.8	142	13	Q42289	Q42289 protein, Homo
21	142.5	30.8	142	13	Q79824	Q79824 protein, Homo
22	142.5	30.8	142	13	Q57436	Q57436 protein, Homo
23	142.5	30.8	142	13	Q93527	Q93527 protein, Homo
24	142	30.5	207	13	Q90X00	Q90X00 protein, Homo
25	140.5	30.3	142	13	Q8AV14	Q8AV14 protein, Homo
26	138.5	29.6	142	13	Q93380	Q93380 protein, Homo
27	137	29.6	142	13	Q93380	Q93380 protein, Homo
28	137	29.6	142	13	Q93380	Q93380 protein, Homo
29	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
30	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
31	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
32	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
33	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
34	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
35	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
36	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
37	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
38	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
39	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
40	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
41	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
42	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
43	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
44	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo
45	136.5	29.5	142	13	Q91A40	Q91A40 protein, Homo

ALIGNMENTS

RESULT 1:

ID	Q8HXV2	PRELIMINARY:	PRE:	110 AA.
AC	Q8HXV2:			
DE	Q1-MAR-2003 (Trembl)	23, Created		
DE	Q1-MAR-2003 (Trembl)	23, Last sequence update		
DE	Q1-MAR-2003 (Trembl)	23, Last annotation update		
DE	Insulin precursor			
GN	INS			
OS	Pongo pygmaeus (Orangutan)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo			
OX	NCBI_TaxID=9606			
AN	111			
RP	SEQUENCE FROM N.A.			
FA	Stead J.D.H., Jeffreys A.J.:			
FF	"Haplotype diversity at the insulin region."			
RG	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY37503; ANO66937.1			
SC	SEQUENCE 110 AA; 12038 MW; 22D2B32B94F520F8 CRC64;			

Query Match 100.0%; Score 463; DB 6; Length 110;
Host Local Similarity 100.0%; Pred. No. 3.9e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	FVACHLGSHLVPAALVCGERFFVPTKTRRAEDIQVGVGLGGGPGAGSLQPLALEG 60
DB	25	FVACHLGSHLVPAALVCGERFFVPTKTRRAEDIQVGVGLGGGPGAGSLQPLALEG 84
QY	61	SLOKRGIVEOCCTSSLSYLENYCN 86
DB	85	SLOKRGIVEOCCTSSLSYLENYCN 110

RESULT 2

ID	Q91X13	PRELIMINARY:	PRE:	110 AA.
AC	Q91X13:			
DE	Q1-DEC-2001 (Trembl)	19, Created		

```

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 20, Last annotation update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sotulidae;
OC Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas.
RA Tredrea M.M., Buck M.J., Guhaniyogi J., Squire T.L., Andrews M.T.;
RT "Regulation of PDK4 expression in a hibernating mammal."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EURL: AK03664; AAK7258.1.
DR HSSP: P01368; IINF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00076; IIGF.
DR PROSITE: PS00262; INSULIN.
SQ SEQUENCE 110 AA: 12069 MW: 4511769D6622BEE5 CRC64:
Query Match: 89.9%; Score 419; DB 11; Length 110;
Best Local Similarity 89.9%; Pred. No. 3; 4e-41;
Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 60
DE 25 FVNHLCGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 44
QY 61 SLQKRGIVCECTSTICSLVQLENYCN 86
DE 85 PLQKRGIVCECTSTICSLVQLENYCN 110

RESULT 3
Q8WNW6 PRELIMINARY; PRT; 119 AA.
AC Q8WNW6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 21, Last annotation update)
DE Preproinsulin.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas.
RA Okamoto S., Marimatsu M.;
RT "cat. insulin."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DE EURL: AB043535; BAB8410.1.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00076; IIGF.
DR PROSITE: PS00262; INSULIN.
SQ SEQUENCE 110 AA: 12069 MW: 95PB6E170C75E7A4 CRC64:
Query Match: 83.9%; Score 388; DB 6; Length 110;
Best Local Similarity 83.9%; Pred. No. 3; 1e-36;
Matches 72; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 60
DB 25 FVNHLCGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 54

```

```

QY 61 SLQKRGIVCECTSTICSLVQLENYCN 86
DB 85 PLQKRGIVCECTSTICSLVQLENYCN 110

RESULT 4
Q8HZ80 PRELIMINARY; PRT; 65 AA.
AC Q8HZ80;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Gorilla gorilla (Gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Sullivan C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EURL: AY092023; AAM76640.1.
DR NON_TER 1.
DR NON_TER 65.
SQ SEQUENCE 65 AA: 6920 MW: B772017FD8BCABEA CRC64:
Query Match: 73.9%; Score 342; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 4; 7e-33;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 66
DB 1 CGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 60

QY 67 IVEQC 71
DB 61 IVEQC 65

RESULT 5
Q8HZ80 PRELIMINARY; PRT; 65 AA.
AC Q8HZ80;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Longo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Sullivan C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EURL: AY092024; AAM76641.1.
DR NON_TER 1.
DR NON_TER 65.
SQ SEQUENCE 65 AA: 6920 MW: B772017FD8BCABEA CRC64:
Query Match: 73.9%; Score 342; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 4; 7e-33;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 66
DB 1 CGSHLVCEALVLCGERGFFYTPKTRREAEDLVQGVELGGPGAGSLQPLALEGSLQKRG 60

```

```

QY 67 IVQC 71
DB 61 IVQC 65

RESULT 6
Q918Q7 PRELIMINARY: PRT: 105 AA.
AC Q918Q7
DT 01-OCT-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20362507; PubMed=10418274;
RA Irwin D.M., Sivarajah P.;
RT "Proinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
RT proinsulin processing."
RL Comp. Biochem. Physiol. 125B:405-410(2000).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF227187; AAF87285.1;
DR HSSP: P01315; ISCB.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin.
DR PRINTS: PR00277; INSULINB.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN. 1.
SQ SEQUENCE 105 AA; 12183 MW; 3A870EEC70217F52 CRC64;

Query Match 53.2%; Score 246.5; DB 13; Length 106;
Best Local Similarity 51.5%; Pred. No. 1.7e-21;
Matches 52; Conservative 5; Mismatches 7; Indels 33; Gaps 4;

QY 1 FVNHLCGSHLVVVALYVCGERGFYTPKTRREAEDELQVQVGLGGPGAGSIQPLA--AL 58
DB 24 FDNQYLCGSHLVVVALYVCGDRGFYSPKSRREAEDELQVQVGLGGPGAGSIQPLA 55
|||||
QY 59 EGS-----EGSR--GIVEQCCTSIQSLYQLENYCN 86
DB 56 QGSELDEMVOVSQAFQKRGIVEQCCHTCSLYDIENYCN 106
|||||

RESULT 7
Q981A7 PRELIMINARY: PRT: 111 AA.
AC Q981A7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (fragment).
OS Osteoglossus bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21203577; PubMed=11306171;
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.R.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199588; AAK28712.1;
DR HSSP: P01308; IHIS.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin. 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN. 1.
SQ SEQUENCE 110 AA; 12324 MW; ADFCCD659D872E06 CRC64;

Query Match 49.9%; Score 230.5; DB 13; Length 110;
Best Local Similarity 46.4%; Pred. No. 1.4e-19;
Matches 45; Conservative 14; Mismatches 13; Indels 25; Gaps 3;

QY 3 NQHLCGSHLVVVALYVCGERGFYTPKTRREAEDELQVQVGLGGPGAGSIQPLA---- 58
DB 26 SQHLCGSHLVVVALYVWCGRGFFYPKTRKRDV-----PLIGFLSPKSAQENE 73
|||||

QY 59 -----EGSLQ--KRGIVEQCCTSIQSLYQLENYCN 86
DB 74 ADEYPFKDGGD-LKVRGIVEQCCHTFCNIFDQNYCN 110
|||||

RESULT 9
Q902Y1 PRELIMINARY: PRT: 110 AA.
AC Q902Y1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (fragment).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Catia.
 CC NCBI_TaxID=72446;
 FN (1)
 RP SEQUENCE FROM N.A.
 RA Bhattacharya S., Roy S.S., Dasgupta S., Ravikumar L., Mukherjee M.,
 RA Bandyopadhyaya I., Wakabayashi K.;
 RT "A new cell secreting insulin";

```

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC EMBL: AF373021; AAK51558.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN 1.
SQ SEQUENCE 108 AA; 11861 MW; D713024E22E95D59 CRC64;

Query Match 45.9%; Score 210.5; DP 13; Length 108;
Best Local Similarity 44.8%; Pred. No. 1.3e-16;
Matches 43; Conservative 12; Mismatches 14; Indels 2; Gaps 0;

QY 4 QHLGSHLVEAFLVCGEGFFYPTKTRKREDFLQVGVGGGAGSLQFLALEGS- 61
DB 27 QHLGSHLVEAFLVCGEGFFYPTKTRKREDFLQVGVGGGAGSLQFLALEGS- 61
QY 62 LQKRGIVEQCCTSTCSLYQLENYCN 86
DB 73 ADFAFKDAEVLKRGVVEGCHKPKCSIFELQNYCN 108

RESULT 13
Q98TA9 PRELIMINARY: PRT; 57 AA.
ID Q98TA9;
AC Q98TA9; (TRENBLrel. 17, Created)
CT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE Preproinsulin (Fragment).
OS Gnathemus petersi.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathemus.
CX NCBI_TaxID=42645;
RN [1]
RP MEDLINE=2103577; PubMed=11306171;
RX Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RE Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN 1.
FT NON_TER 87
FT NON_TER 87
SQ SEQUENCE 87 AA; 9874 MW; FF446ED35D245385 CRC64;

Query Match 45.9%; Score 210.5; DP 13; Length 87;
Best Local Similarity 45.6%; Pred. No. 2.4e-17;
Matches 43; Conservative 11; Mismatches 28; Indels 3; Gaps 2;

QY 4 QHLGSHLVEAFLVCGEGFFYPTKTRKREDFLQVGVGGGAGSLQFLALEGS- 61
DB 4 QHLGSHLVEAFLVCGEGFFYPTKTRKREDFLQVGVGGGAGSLQFLALEGS- 61
QY 62 LQKRGIVEQCCTSTCSLYQLENYCN 86
DB 63 KVRKRGIVEQCCHPCNFDLQNYCN 87

RESULT 14
Q98TB1 PRELIMINARY: PRT; 108 AA.
ID Q98TB1;
AC Q98TB1; (TRENBLrel. 17, Created)
CT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE Preproinsulin (Fragment).
OS Ambloplites rupestris (Rock bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Centrarchidae; Ambloplites.
CX NCBI_TaxID=109273;
RN [1]
RP MEDLINE=2103577; PubMed=11306171;
RX Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNA from the rock bass."
RT Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN 1.
FT NON_TER 1
FT NON_TER 91
SQ SEQUENCE 91 AA; 10100 MW; E86C8B256DC69D39 CRC64;

Query Match 44.0%; Score 203.5; DP 13; Length 91;

```

```

Best Local Similarity 46.7%  Prod. No. 1.8e-10
Matches 42: Conservative 13: Mismatches 26: Indels 5: Gaps 4:
QY 4 QHLCGSHLVKALYIVCGSRGFFYTKTRRFRADIC--VSLVELGGSPGA-GSHLDGALD 89
   ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4 QHLCGSHLVKALYIVCGSRGFFYTKTRRFRADIC--VSLVELGGSPGA-GSHLDGALD 89
   ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 GSTCQ---KRGIVECCCTSGCSLVQIERYCN 86
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 DQEMMYKRGIVECCCHHPNCFDLSRYCN 91
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: September 15, 2003, 12:03:29
Job time : 35.5986 secs

OM protein - protein search, using sw model
 Run on: September 15, 2003, 11:44:15 : Search time 46.9326 seconds
 (without alignments)
 225,942 million cell updates/sec

Title: US-09-423-100-5
 Perfect score: 294
 Sequence: 1 FVHQELQGSHEVFLVIVVQ3.....FVEQCTSGSDVGLNYSN 32
 Scoring table: BLCSUM62
 Gapop 10.0 / Gapext 0.5
 Searched: 1107653 seqs, 15672673 residues
 Total number of hits satisfying chosen parameters: 1137863
 Minimum DB seq length: 8
 Maximum DB seq length: 200000000
 Post-processing: Minimum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_19Jur03:*

1:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1580.DAT:
2:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:
3:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:
4:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:
5:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:
6:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:
7:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:
8:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:
9:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:
10:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:
11:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:
12:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:
13:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:
14:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:
15:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:
16:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:
17:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:
18:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:
19:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:
20:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:
21:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:
22:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:
23:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:
24:	/SDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY:

Result No.	Query Match	Length DB	ID	Description
1	294	100.0	52	20 AAY42859 Human insulin precursor
2	294	100.0	56	15 AAR71695 Human mature insulin
3	294	100.0	56	15 AAR78665 Human proinsulin
4	294	100.0	63	15 AAR8900 Human proinsulin
5	294	100.0	96	15 AAR8899 Human proinsulin
6	294	100.0	96	15 AAR78662 Human proinsulin
7	294	100.0	107	20 AAY42860 Fused protein
8	294	100.0	116	17 AAR8867 hGH mature proinsulin
9	294	100.0	137	15 AAR71692 Mating factor alpha

10	294	100.0	145	16 AAK71694 Mating factor alpha
11	294	100.0	146	16 AAR71695 Mating factor alpha
12	294	100.0	150	20 AAY42861 Chimeric protein
13	294	99.0	57	11 AAR4582 Proinsulin analog
14	294	97.6	52	12 AAR11899 Example of human I
15	293.5	96.4	53	15 AAR55883 Di-Arg-(B31-32)-Hu
16	293.5	96.4	53	18 AAM18007 Ins1 double-chain
17	293.5	96.4	160	16 AAR79056 Glycosylphosphatid
18	292.5	96.1	127	17 AAR98896 SOD-proinsulin hyb
19	291.5	95.7	60	3 AAP20002 Human proinsulin a
20	291.5	94.7	51	21 AAB1277 Human mature insul
21	291.5	94.7	51	21 AAB1277 Human mature insul
22	291.5	94.7	51	23 AAM48219 Human insulin, Ro
23	291.5	94.7	53	17 AAR96046 Modified pJ559 mod
24	291.5	94.7	55	17 AAR6050 Modified pJ59 mod
25	291.5	94.7	63	17 AAM0807 Single chain insul
26	291.5	94.7	66	17 AAM0809 Met-Arg-Met-single
27	291.5	94.7	87	17 AAM0810 Single chain insul
28	291.5	94.7	138	17 AAR7086 PKV142 modified in
29	291.5	94.4	53	7 AAP0132 Sequence of novel
30	291.5	94.4	53	15 AAR65882 Gly-A21-di-Arg (B3
31	291.5	94.2	65	19 AAM4765 Preproinsulin J.
32	291.5	94.0	67	18 AAM17999 Ins1 double-chain
33	291.5	94.0	67	13 AAM17998 Ins2 double-chain
34	291.5	93.7	53	12 AAR11898 Example of human I
35	291.5	93.7	53	17 AAR7085 pAK679 modified in
36	291.5	93.7	53	20 AAY28341 Modified Human ins
37	291.5	93.7	53	22 AAR30704 Amino acid sequenc
38	291.5	93.7	55	8 AAP71019 Sequence encoded b
39	291.5	93.7	58	17 AAR96047 Modified pJ559 rod
40	291.5	93.7	59	17 AAR96048 Modified pJ559 mod
41	291.5	93.7	61	17 AAR96049 Modified pJ559 mod
42	291.5	93.7	65	17 AAR8188 N-terminally exten
43	291.5	93.7	89	17 AAR8179 Signal peptide/lea
44	291.5	93.7	91	13 AAR25584 Insulin precursor
45	291.5	93.7	91	17 AAR8181 YAP3 signal peptid

ALIGNMENTS

RESULT 1
 AAY42859
 1-2 AAY42859 standard; protein; 52 AA.
 XX AAY42859;
 AC AAY42859;
 XX :9-JAN-2000 (first entry)
 XX Human insulin precursor, SEQ ID 5.
 XX
 XX Insulin: precursor; growth hormone; chaperone; intramolecular;
 XX folding; conformation; chimeric protein; cleavable; recombinant;
 XX production; yield.
 XX Homo sapiens.
 XX W09550302-AL.
 XX 07-OCT-1999.
 XX 31-MAR-1998; 98WC-CN000052.
 XX 31-MAR-1998; 98WC-CN000052.
 XX (TONGSHI) TONGSHI GANTECH BIOTECHNOLOGY LTD.
 XX Gen 2;
 XX WPI: 1999-610839/52.
 XX New chimeric proteins containing human growth hormone fragment, used
 XX particularly for the production of human insulin -
 PT

XX PS Claim 12; Page 29 30; 46pp; English.

XX CC This sequence represents a human insulin precursor comprising insulin A and B chains. This insulin precursor is a component of the chimeric proteins BGM-mini-proinsulin (AA42866) and the chimeric protein given in AA42867. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a cleavable peptide linker (AA42857). The hGH portion of the chimeric protein acts as an intramolecular chaperone (23C) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulfitolysis and related steps of purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.

XX SQ Sequence 52 AA:

Query Match 100.0% Score 294; DB 29; Length 52;
 Best Local Similarity 100.0% Pred. No. 1.7e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNHLCGSHLVFALVLCVGERGFFYTPKTRGVVEOCCTSGSLVQLENYCN 52
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 5 FVNHLCGSHLVFALVLCVGERGFFYTPKTRGVVEOCCTSGSLVQLENYCN 52

RESULT 2

AA68901

ID AAR68901 standard; peptide: 56 AA.

XX AC AAR68901;

XX XX 25-MAR-2003 (updated)

XX DI 02-MAR-1995 (first entry)

XX XX Human pro-insulin 3.

XX DE Pro-insulin; A-chain; B-chain; C-chain; disulphide;

XX KW mercaptan; chaotropic agent.

XX OS Homo sapiens.

XX XX EP600372-A1.

XX XX 08-JUN-1994.

XX XX 25-NOV-1993; 93EP-0118993.

XX PR 02-DEC-1992; 92DE-4240420.

XX XX (FARH) HOECHST AG.

XX XX Gerl M, Ludwig J, Obermeier S, Sabel W;

XX XX WPI; 1994-177718/22.

XX DR Prodn. of pro-insulin with correct disulphide bridges - by

XX PT treating recombinant precursor protein with mercaptan - by

XX PT and in presence of chaotropic agent, then isolation on

XX PT hydrophobic resin

XX PS Disclosure; Page 12; 15pp; German.

XX CC Proinsulin is produced by treating recombinant precursor protein with a mercaptan to provide 2-10 SH residues per Cys residue, in presence of a chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3-50 g hydrophobic adsorber resin per 1 aq. medium of pH 4-7, isolating the adsorbed resin and pro-insulin and desorbing the pro-insulin. This method produces pro-insulin with correctly bonded Cys bridges. Compared with known methods it involves fewer stages (esp. no sulfitolysis or cyanogen bromide cleavage) and overall losses during purification are reduced, i.e. the process is quicker and gives better yields.

XX CC Sequences of insulin chain A, B and C are given in AAR68895-97.

XX CC Sequences of pro-insulin 1-4 are given in AAR68898-901.

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 56 AA:

Query Match 100.0% Score 294; DB 15; Length 56;
 Best Local Similarity 100.0% Pred. No. 1.7e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNHLCGSHLVFALVLCVGERGFFYTPKTRGVVEOCCTSGSLVQLENYCN 52
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 5 FVNHLCGSHLVFALVLCVGERGFFYTPKTRGVVEOCCTSGSLVQLENYCN 56

RESULT 3

AA68665

ID AAR68665 standard; protein: 56 AA.

XX XX AAR68665;

XX XX 03-APR-1996 (first entry)

XX XX Proinsulin sequence 3.

XX XX Proinsulin; post-translational modification; recombinant production;

XX KW protein folding; conformation.

XX XX Synthetic.

XX XX Key Location/Qualifiers

XX FT Region 1..4

XX FT /label= R2

XX FT /note= "a peptide of 4 amino acids"

XX FT Peptide 5..34

XX FT /label= R1-(B2-B29)-Y

XX FT Region 35

XX FT /note= "human insulin B-chain"

XX FT Peptide 36..56

XX FT /label= X

XX FT /label= Gly-(A2-A20)-R3

XX FT /note= "human insulin A-chain"

XX XX EF668292-A2.

XX XX 23-AUG-1995.

XX XX 09-FEB-1995; 95EP-01C1746.

XX XX 18-FEB-1994; 94DE-4405179.

XX XX (FARH) HOECHST AG.

XX XX Gerl M, Ludwig J, Obermeier R, Sabel W;

XX XX WPI; 1995-284754/38.

XX XX Isolation of insulin that is correctly post-translationally

XX PT processed - by reacting proinsulin with a mercaptan in the presence

XX PT of a chaotropic agent and purificn. after absorption to hydrophobic resin

XX XX

pro-insulin is produced by treating recombinant precursor pro-insulin with a mercaptan to provide 2-10 SH residues per Cys residue, the presence of a chaotropic agent and in an aq. medium of pH 10-11, treating the prod. with 3-50 g hydrophobic adsorber resin per l. and medium of pH 4-7, isolating the adsorbed resin and pro-insulin and desorbing the pro-insulin. This method produces pro-insulin with correctly bonded Cys bridges. Compared with known methods it

RESULT 9
AAR71692

[illegible]

DT	20-NOV-1995	{first entry}	
xx			
DE			Mating factor alpha 1-Insulin precursor Arg81, Arg831 N-terminal.


```

XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex
KW mating factor alpha 1; N-terminal BEAEAEAR.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Protein: 1..85
FT Peptide: /label= mating factor alpha 1
FT Peptide: 86..93
FT Peptide: /label= N-terminal peptide
FT Peptide: 94..124
FT Peptide: /label= B-chain
FT Peptide: 125..145
FT Peptide: /label= A-chain
XX
XX WO9507931-A1.
XX 23-MAR-1995.
XX 16-SEP-1994; 94WO-DK00347.
XX 17-SEP-1993; 93DK-0001044.
XX 02-FEB-1994; 94US-0190829.
XX (NOVO ) NOVO-NORDISK AS.
XX Andersen AS, Halstrom JB, Havelund S, Jonassen J;
PI Markussen J;
XX WPI: 1995-131314/17.
XX N-PSDB; AAQ86423.
XX Acylated insulin deriv. which may be present as a zinc ion
XX complex - is used to treat diabetes and is rapid acting.
XX Example 5: Pages 82-83; 100pp; English.
XX
XX AAQ86429 encodes AAR71694 mating factor alpha 1-insulin precursor
CC ArgB1, ArgB31 N-terminal BEAEAEAR. The insulin precursor comprises
CC the B and A chains of a claimed human insulin derivative preceded
CC by the N-terminal amino acids BEAEAEAR. In the final claimed compsn.
CC they are covalently connected via disulphide bonds between Cys
CC residues A7/B7 and A20/B19. The derivative, which may be present
CC as a zinc ion complex, can be used as a fast action treatment for
CC diabetes.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 145 AA:
XX
XX Query Match 100.0%; Score 294; DB 16; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-26;
XX Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 FVNQHLCGSHLVEAALYVGGERGFYTPKTRGVQCCCTSGSYOLENYCN 52
111 1111 1111 111 1111 1111 111 111 111 1
DB 94 FVNQHLCGSHLVEAALYVGGERGFYTPKTRGVQCCCTSGSYOLENYCN 145
111 1111 1111 111 1111 1111 111 111 111 1
RESULT 11
AAR71695
ID AAR71695 standard; Protein: 146 AA.
XX
XX AAR71695:
XX AC
XX 25-MAR-2003 (updated)
XX 20-NOV-1995 (first entry)
XX
XX Mating factor alpha 1-insulin precursor ArgB1, ArgB31 N-terminal.
XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
KW mating factor alpha 1; N-terminal BEAEAEAR.
XX

```

```

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Protein: 1..85
FT Peptide: /label= mating factor alpha 1
FT Peptide: 94..94
FT Peptide: /label= N-terminal peptide
FT Peptide: 95..125
FT Peptide: /label= B-chain
FT Peptide: 126..146
FT Peptide: /label= A-chain
XX
XX WO9507931-A1.
XX 23-MAR-1995.
XX 16-SEP-1994; 94WO-DK00347.
XX 17-SEP-1993; 93DK-0001044.
XX 02-FEB-1994; 94US-0190829.
XX (NOVO ) NOVO-NORDISK AS.
XX Andersen AS, Halstrom JB, Havelund S, Jonassen J;
PI Markussen J;
XX WPI: 1995-131314/17.
XX N-PSDB; AAQ86432.
XX Acylated insulin deriv. which may be present as a zinc ion
XX complex - is used to treat diabetes and is rapid acting.
XX Example 6: Page 85; 100pp; English.
XX
XX AAQ86432 encodes AAR71695 mating factor alpha 1-insulin precursor
CC ArgB1, ArgB31 N-terminal BEAEAEAR. The insulin precursor comprises
CC the B and A chains of a claimed human insulin derivative preceded
CC by the N-terminal amino acids BEAEAEAR. In the final claimed compsn.
CC they are covalently connected via disulphide bonds between Cys
CC residues A7/B7 and A20/B19. The derivative, which may be present
CC as a zinc ion complex, can be used as a fast action treatment for
CC diabetes.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 146 AA:
XX
XX Query Match 100.0%; Score 294; DB 16; Length 146;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-26;
XX Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 FVNQHLCGSHLVEAALYVGGERGFYTPKTRGVQCCCTSGSYOLENYCN 52
111 1111 1111 111 1111 1111 111 111 111 1
DB 95 FVNQHLCGSHLVEAALYVGGERGFYTPKTRGVQCCCTSGSYOLENYCN 146
111 1111 1111 111 1111 1111 111 111 111 1
RESULT 12
AAR72861
ID AAY42861 standard; protein: 150 AA.
XX
XX AAY42861:
XX AC
XX 19-JAN-2000 (first entry)
XX
XX Chimeric protein, SEQ ID 7.
XX
XX Insulin; precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO950302-A1.
XX

```

XX 07-OCT-1999.
 XX
 XX 31-MAR-1998: 98WO-CR00052.
 XX
 XX 31-MAR-1998: 98WO-CR00052.
 XX
 XX (TONG:) TONGHUA GANTHEM BIOTECHNOLOGY LTD.
 XX
 XX Gan Z:
 XX
 XX WPI: 1999 610839/52.
 XX
 XX New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
 XX
 XX Claim 14: Page 30-31: 46pp: English.
 XX
 XX This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AA42856, a cleavable peptide linker (AA42857), and a human insulin precursor comprising insulin A and B chains (AA42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place.
 XX Production of recombinant human insulin via an hGH-proinsulin chimera protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphatolysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.
 XX
 XX Sequence 150 AA:
 SQ
 Query Match 100.0%; Score 294; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 4.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
 DB 99 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 150
 RESULT 13
 AAR04582
 ID AAR04582 standard; protein: 57 AA.
 XX
 XX AAR04582;
 XX
 XX 25-MAR-2003 (updated)
 XX 14-SEP-1990 (first entry)
 XX
 XX Proinsulin analogue with a Lys residue linking the A and B chains.
 XX
 XX insulin fusion protein; pre-insulin analogue; tandemistore;
 KW Lys-Lys bridge: ds
 XX
 XX synthetic.
 OS
 XX Key Location/Qualifiers
 PH misc-difference 36
 FT /label-Lys residue linking insulin B chain to A chain
 FT Peptide 1..35
 FT /label- Insulin B chain
 FT Peptide 37..57

FT
 XX EP367183-A.
 XX
 XX 09-MAY-1990.
 XX
 XX 22-OCT-1989: 89EP-0120056.
 XX
 XX 03-NOV-1988: 88CE-3837273.
 XX 19-AUG-1989: 89DE-3927449.
 XX
 XX (FARH) ROECHST AG.
 XX
 XX Kolter KP, Riess GJ, Uhlmann E, Wallmeier H;
 XX WPI: 1990-141149/19.
 XX N-PSDB: AAQ04335.
 XX
 XX New insulin fusion proteins comprise pro-insulin analogue linked to tandemistate
 XX
 XX Disclosure: Page 2: 2pp: Gorman.
 XX
 XX This sequence is joined to the C-terminus of an N-terminal fragment comprising opt. modified tandemistate. This fusion protein may be converted into human insulin using known methods. The synthetic gene was prepared by the phosphoramidite method.
 XX See also AAQ04335.
 XX (Updated on 25-MAR-2003 to correct PR field.)
 XX (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 57 AA:
 SQ
 Query Match 99.08; Score 291; DB 11; Length 57;
 Best Local Similarity 98.1%; Pred. No. 3.8e-26;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
 DB 6 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 57
 RESULT 14
 AAR11899
 ID AAR11899 standard; Protein: 52 AA.
 XX
 XX AAR11899;
 XX
 XX 25-MAR-2003 (updated)
 XX 22-JUL-1991 (first entry)
 XX
 XX Example of human insulin precursor.
 XX
 XX Human insulin; diabetes; transpeptidation.
 XX
 XX BORO suppliers.
 XX
 XX EP427296-A.
 XX
 XX 15-MAY-1991.
 XX
 XX 29-MAY-1985: 90EP-0121887.
 XX
 XX 15-NOV-1990: 90SP-0121887.
 XX 30-MAY-1984: 84DK-0002665.
 XX 08-FEB-1985: 85DK-0005082.
 XX
 XX (NWOV) NOVVO-NORDISK AS.
 XX
 XX Markussen J, Fjell N, Ammerer G, Hansen MT, Thim L;
 XX Norris K. Volgt 80;
 XX WPI: 1991-141828/20.

/label- Insulin A chain

```

XX human insulin precursors - expressed with correctly positioned
PT disulphide bridges giving improved resistance to proteolysis
XX
XX Claim 3: page 18: 28pp: English.
XX
CC This human insulin precursor has correctly positioned disulphide
CC bridges between the A and B chains and is more resistant to
CC proteolytic digestion than prior art insulin precursors. Yeast
CC strains transformed with DNA encoding this precursor can be
CC cultured to secrete it in high yields. The precursor can be
CC converted into mature human insulin by transpeptidation.
CC See also AAR11897-98.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 52 AA:
      Query Match          97.6%; Score 287; DB 12; Length 52;
      Best Local Similarity 95.2%; Pred. No. 9, 5e-26;
      Matches 50; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
      1 FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 52
      : 1 FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 52
      : FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 52

RESULT 15
AAR65883
ID AAR65883 standard: protein; 53 AA.
XX
XX AAR65883;
XX
XX 25-MAR-2003 (updated)
XX 26-JUN-1995 (first entry)
XX
XX Di-Arg-(831-32)-human insulin amorphous, monospherical deriva.
XX Human insulin; recombinant production; amorphous; monospherical form;
XX diabetes mellitus.
XX Homo sapiens (produced recombinantly in Escherichia coli).
XX
XX Key Location/Qualifiers
XX Protein 1..30
XX /label= insulin_B-chain
XX Protein 33..53
XX /label= insulin_A chain
XX
XX EP622376-A1.
XX
XX 02-NOV-1994.
XX
XX 27-APR-1994; 94EP-0106196.
XX
XX 27-APR-1993; 93BE-4313702.
XX
XX (FARH ) HCECHST AG.
XX
XX Deil P, Geisen X, Obermeier R, Sabel W;
XX WPI: 1994-334579/42.
XX
XX Amorphous, mono-spherical form of insulin deriva. - for treating
XX diabetes mellitus, are produced by diluting soln. in aq.
XX isopropanol, are stable when dried or in suspension
XX
XX Example 2: Page 5: 10pp: German.
XX
XX This sequence is a specific example of an insulin derivative which
XX can be obtained in amorphous, monospherical form by dissolving in
XX an n-propanol/buffer mixture (pH 4.5-6.5) having n-propanol content
XX 15% relative to water. The solution is then diluted with water to

```

```

CC reduce n-propanol content to below 15%. The resulting insulin
CC preparation is stable and can be used for the treatment of diabetes
CC mellitus.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 53 AA:
      Query Match          96.4%; Score 283.5; DB 15; Length 53;
      Best Local Similarity 98.1%; Pred. No. 2.5e-25;
      Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
      1 FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 52
      : 1 FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 53
      : FVNHLCGSHLVEALYLVGCGRGGFYTPKTRGIVGQCCSTCSLYOLENYCN 53

Search completed: September 15, 2003, 12:00:59
Job time : 37.5305 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003 11:57:40 : Search time 11.9243 seconds
(without alignments)
164.449 Million cell updates/sec

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 FVNHGSSHLVERALVCG.....LVKCCISLGSLEYHENYK 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 3.5

Searched: 328717 seqs, 4210658 residues

Total number of hits satisfying chosen parameters: 42617

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 1%
Maximum Match 100%
Listing first 15 summaries

Database : Issued_Patents_AA*
1: /cgn2_6/prodata/1/iaa/5A.COMB.pept*
2: /cgn2_6/prodata/1/iaa/5B.COMB.pept*
3: /cgn2_6/prodata/1/iaa/6A.COMB.pept*
4: /cgn2_6/prodata/1/iaa/6B.COMB.pept*
5: /cgn2_6/prodata/1/iaa/6C.COMB.pept*
6: /cgn2_6/prodata/1/iaa/6D.COMB.pept*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description
1	294	100.0	56	1	US-08-160-376A-7
2	294	100.0	56	1	US-08-389-487-11
3	294	100.0	63	1	US-08-160-376A-6
4	294	100.0	66	1	US-08-291-040B-5
5	294	100.0	96	1	US-08-160-376A-5
6	294	100.0	96	1	US-08-389-487-8
7	294	100.0	137	1	US-08-400-256-39
8	294	100.0	137	3	US-08-975-365-39
9	294	100.0	145	1	US-08-400-256-45
10	294	100.0	145	3	US-08-975-365-45
11	294	100.0	146	1	US-08-400-256-48
12	294	100.0	146	3	US-08-975-365-48
13	291	99.0	57	1	US-08-000-732A-44
14	293.5	96.4	53	1	US-08-273-617-4
15	293.5	96.4	53	4	US-08-981-988A-42
16	278.5	94.7	51	4	US-09-477-924-3
17	278.5	94.7	51	4	US-09-723-891-3
18	278.5	94.7	51	4	US-09-723-896-3
19	277.5	94.4	53	1	US-08-233-617-3
20	277	94.2	65	3	US-08-900-574-3
21	276.5	94.0	55	3	US-08-900-574-6
22	276.5	94.0	66	3	US-08-900-574-5
23	276.5	94.0	67	4	US-08-981-988A-1
24	276.5	94.0	67	4	US-08-981-988A-5
25	276	93.9	67	3	US-08-900-574-7
26	275.5	93.7	53	3	US-09-261-853-2
27	275.5	93.7	65	1	US-08-468-674B-71

Sequence 71, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 45, Appl
Sequence 45, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 37, Appl
Sequence 3, Appl
Sequence 36, Appl
Sequence 19, Appl
Sequence 33, Appl
Sequence 42, Appl
Sequence 33, Appl
Sequence 42, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-160-376A-7
Sequence 7, Application US/08150376A
Patent No. 5475049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Ranier
APPLICANT: Geil, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cystine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Geroni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Mauser, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/POCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-7

Query Match 100.0%; Score 294; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.9e-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: PatentIn Release #1.0, Version #1.25
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/08/400,256
7 FILING DATE: 03-MAR-1995
8 ATTORNEY/AGENT INFORMATION:
9 CLASSIFICATION: 514
10 NAME: Lambiris, Elias J.
11 REGISTRATION NUMBER: 33,728
12 REFERENCE/DOCKET NUMBER: 3985,220-US
13 TELECOMMUNICATION INFORMATION:
14 TELEPHONE: 212-867-0123
15 TELEFAX: 212-878-9655
16 INFORMATION FOR SEQ ID NO: 48:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 146 amino acids
19 TYPE: amino acid
20 TOPOLOGY: linear
21 MOLECULE TYPE: protein
22 US-08-400-256-48
23
24 Query Match 100.0%; Score 294; DB 1; Length 146;
25 best local Similarity 100.0%; Pred. No. 1.4e-28;
26 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
97
```


Db 2 FVNQHLGGSHLVEALYLVCGERGFFYTPKTRGVLEGGCTSLGSLYQLNQYN 53

RESULT: 15

```

US-08-981-985A-42
: Sequence 42, Application US-0981-985A
: Patent No. 5337-94
: GENERAL INFORMATION
: APPLICANT: Vittal Malviya Scientific Research Foundation
: APPLICANT: The University of Leicester
: TITLE OF INVENTION: Insulin
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: VITTAL MALVIYA SCIENTIFIC RESEARCH FOUNDATION
: STREET: K. K. ROAD
: CITY: BANGALORE
: COUNTRY: INDIA
: ZIP: 560 004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.0 (BPC)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US-08/381-985A

```

Query Match 96.4% Score 283.5 LR 4 Length 52

Best Local Similarity 98.19; Pred. No. 9.80-29;

Best local similarity: 50.14, frequency: 0.002,
Matches 52; Conservative 0; Mismatches 0; Index 1; Gaps 1

1 FVNQHLGSSHLVEALYLVCGERGEFFYPKT-RGTVFQCCTSTCSLVQIENYCK

[illegible]

Search completed: September 15, 2003, 12:05:32
Job time : 11.9283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: September 15, 2003, 12:03:35 : Search time 21.8065 Seconds
(without alignments)
347,945 Million cell updates/sec

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 FVNQRLQGSHEVIAIVYVCG.....IIVEQCSISQNSVNCIENTCN 52

Scoring table: BLOSUM62

Gapop 16.0 , Gapext 6.5

Searched: 54,936 seqs, 1459,2426 residues

Total number of hits satisfying chosen parameters: 54186

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published_Applications_AA:

- 1: /cgn2_6/ptcddata/1/pubaa/us07_PURCOMB.pep:
- 2: /cgn2_6/ptcddata/1/pubaa/PCT_NEW_PUB.pep:
- 3: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep:
- 4: /cgn2_6/ptcddata/1/pubaa/US06_PURCOMB.pep:
- 5: /cgn2_6/ptcddata/1/pubaa/US07_NEW_PUB.pep:
- 6: /cgn2_6/ptcddata/1/pubaa/PCT_NEW_PUB.pep:
- 7: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep:
- 8: /cgn2_6/ptcddata/1/pubaa/US08_PURCOMB.pep:
- 9: /cgn2_6/ptcddata/1/pubaa/US09A_PURCOMB.pep:
- 10: /cgn2_6/ptcddata/1/pubaa/US09A_PURCOMB.pep:
- 11: /cgn2_6/ptcddata/1/pubaa/US09C_PURCOMB.pep:
- 12: /cgn2_6/ptcddata/1/pubaa/US09C_PURCOMB.pep:
- 13: /cgn2_6/ptcddata/1/pubaa/US10A_PURCOMB.pep:
- 14: /cgn2_6/ptcddata/1/pubaa/US10A_PURCOMB.pep:
- 15: /cgn2_6/ptcddata/1/pubaa/US10C_PURCOMB.pep:
- 16: /cgn2_6/ptcddata/1/pubaa/US12_NEW_PUB.pep:
- 17: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep:
- 18: /cgn2_6/ptcddata/1/pubaa/US06_PURCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	52	14	US-10-054-873-5
2	294	100.0	107	14	US-10-054-873-5
3	294	100.0	150	14	US-10-054-873-7
4	278.5	94.7	5	11	US-09-658-935A-5
5	278.5	94.7	5	14	US-10-028-410-3
6	275.5	93.7	124	10	US-09-894-711-13
7	275.5	93.7	138	10	US-09-861-687-19
8	273	92.9	50	14	US-10-068-009A-3
9	267	90.8	86	10	US-09-878-380-1
10	267	90.8	86	11	US-09-858-945A-4
11	267	90.8	86	14	US-10-028-410-2
12	267	90.8	86	14	US-10-054-873-4
13	267	90.8	96	10	US-09-947-563-4
14	267	90.8	110	5	US-09-265-854-12
15	267	90.8	110	5	US-09-915-229-5

16	267	90.8	110	10	US-09-804-409A-9	Sequence 9, Appl1
17	267	90.8	110	12	US-09-969-7480-6	Sequence 6, Appl1
18	267	90.8	110	15	US-10-038-686-1	Sequence 1, Appl1
19	267	90.8	110	15	US-10-328-813-2	Sequence 2, Appl1
20	267	90.8	117	9	US-09-280-030-63	Sequence 63, Appl1
21	267	90.8	130	9	US-09-280-030-62	Sequence 62, Appl1
22	266.5	90.6	124	9	US-09-736-611-12	Sequence 12, Appl1
23	266.5	90.6	124	9	US-09-740-359-12	Sequence 12, Appl1
24	266.5	90.6	124	10	US-09-894-711-12	Sequence 12, Appl1
25	266.5	90.6	125	9	US-09-736-611-10	Sequence 10, Appl1
26	266.5	90.6	125	9	US-09-740-359-10	Sequence 10, Appl1
27	266.5	90.6	125	10	US-09-894-711-10	Sequence 10, Appl1
28	266.5	90.6	147	9	US-09-736-611-8	Sequence 8, Appl1
29	266.5	90.6	147	9	US-09-740-359-7	Sequence 7, Appl1
30	265	90.1	144	9	US-09-736-611-6	Sequence 6, Appl1
31	265	90.1	144	9	US-09-740-359-5	Sequence 5, Appl1
32	265	90.1	146	10	US-09-894-711-5	Sequence 5, Appl1
33	261	88.8	96	10	US-09-947-563-5	Sequence 7, Appl1
34	244.5	83.2	144	10	US-09-894-711-7	Sequence 132, App
35	219	74.5	46	9	US-09-205-658-132	Sequence 133, App
36	214	72.8	46	9	US-09-205-658-133	Sequence 135, App
37	209	68.0	46	9	US-09-205-658-135	Sequence 134, App
38	195	66.3	46	9	US-09-205-658-134	Sequence 136, App
39	194	66.0	46	9	US-09-205-658-137	Sequence 137, App
40	184	62.6	46	9	US-09-205-658-137	Sequence 115, App
41	183.5	62.4	50	9	US-09-815-229-14	Sequence 14, Appl1
42	173	58.8	32	9	US-09-947-563-6	Sequence 6, Appl1
43	173	58.8	32	10	US-09-947-563-6	Sequence 10, Appl1
44	173	58.8	56	9	US-09-815-229-10	Sequence 805, App
45	173	58.8	156	9	US-09-925-297-805	

ALIGNMENTS

RESULT :
US-10-054-873-5
Sequence 5, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew L.P.
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3634
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DCS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/G0052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid

```

; STRANDEDNESS: UNKNOWN
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID:
CS-10-054-B73-K

```

```

Query Match      100.00; Score 294; DB 117; Length 141
Best Local Similarity 100.00; E-Val 2.5E-20
Matches 52; Conservative 0; Mismatch 0; Gaps 0;
      QY 2 PYNHILCSHVAEALVACSGEGFFETKTSIVGQNTCSLSVHENVK 52
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      DB 1 PYNHILCSHVAEALVACSGEGFFETKTSIVGQNTCSLSVHENVK 52
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2
US-10-054-973-f
: Sequence 6, Application US/10054873
: Publication No. US2010167121

GENERAL INFORMATION:
APPLICANT: Gan, Zhong Hu
TITLE OF INVENTION: Chimeric Protein Containing an Intracellular Chapterone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

```

1 COUNTRY: USA
2
3 ZIP: 94111-3834
4
5 COMPUTER READABLE FORM:
6
7 MEDIUM TYPE: Floppy disk
8
9 COMPUTER: IBM PC compatible
10
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12
13 SOFTWARE: Patentir Release #1.0. Version: #1.30
14
15 CURRENT APPLICATION DATA:
16
17 APPLICATION NUMBER: US/10/054,873
18
19 FILING DATE: 22-Jan-2002
20
21 CLASSIFICATION: <unknown>

```

```

1  CLASSIFICATION: <UNCLASS>
2
3  PRIOR APPLICATION DATA:
4
5  APPLICATION NUMBER: WO PCT/CN96/00052
6  FILING DATE: 31-MAR-1998
7  APPLICATION NUMBER: US 09/423,110
8  FILING DATE: 11-DEC-2000
9  ATTORNEY/AGENT INFORMATION:
10 NAME: MYCIGHT, FRANK J.
11 REGISTRATION NUMBER: 46,946
12 REFERENCE/DOCKET NUMBER: 020167-000-0045

```

```

1 REFERENCE/LOCK# NUMBER: 020167-000-3100
2
3 INFORMATION FOR SEQ ID NO: 1:
4     SEQUENCE CHARACTERISTICS:
5         LENGTH: 107 amino acids
6         TYPE: amino acid
7         STRANDEDNESS: <UNKNOWN>
8         TOPOLOGY: linear
9     MOLECULE TYPE: protein
10     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
11     DS-10-054-873-6

```

Query Match	100.00;	Score	294;	LR	14;	Length	107;
Best Local Similarity	100.00;	pred. No.	5.5e-29;				
Matches	52;	Conservative	0;	Mismatches	0;	Indels	0;

Qy 3 FVNHLCGSHLVKALYLACGGSGFFFTKTRIGVEQCCTSCSLYULENYCN 52
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 56 FVNHLCGSHLVKALYLACGGSGFFFTKTRIGVEQCCTSCSLYULENYCN 108

RESULT 3
US-10-054-873-7
; Sequence 7, Application: US/10054873
; Publication No. US20020164712A;
; GENERAL INFORMATION;

APPLICANT: Gao, Zhong Ku
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Crew LLP
STREET: 170 Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-2834

COMPUTER READABLE FORM:
ZIP: 94111-5634
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,673
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknow>

PRIOR APPLICATION DATA: WO PCT/CN98/00052
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 13-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 45,946
REFERENCE/DOCKET NUMBER: 020167-00013005

SEQUENCE CHARACTERISTICS:
 LENGTH: 150 amino acids
 TYPE: amino acid
 STRANDEFNS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 4-873-7

Match:	100.0%	Score 294;	DB 14;	Length 150;
% Identical Similarity	100.0%	Pred. NO. 7.9e-29;		
s 52: Conservative	0;	Mismatches	0;	Indels 0

```

1 FVNHLCSSH:VEALYLVCGGEGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 52
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 FVNHLCSSH:VEALYLVCGGEGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 150

```

4-935B-5
Page 5, Application US/69658935B
Application No. US2003005917A1
4. INFORMATION:
INVENTOR: Dubaque, Yves
INVENTOR: Flivartoff, Ellen
INVENTOR: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
REFERENCE: P1794*

REFERENCE: PI794E.
 INST APPLICATION NUMBER: US09/858,335B
 FILING DATE: 2002-07-02
 APPLICATION NUMBER: US 60/248,985
 FILING DATE: 2000-11-15
 APPLICATION NUMBER: US 60/204,490
 FILING DATE: 2000-05-16
 NO. OF SEQ ID NOS: 153

NO 3
TH: 51
: PRF
NTSM: Homo sapiens
8-935H-5

Match	94.7%	Score 278.5;	DB 11;	Length 51;
Local Similarity	98.1%	Pred. No. 2e-27;		

Matches 51: Conservative 0: Mismatches 0: Indels 1: Gaps 1:
QY : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 51

RESULT 5
US-10-028-410-3
: Sequence 3, Application US/1028410
: Publication No. US20020160955A1
: GENERAL INFORMATION:
: APPLICANT: Debaquie, Yves
: APPLICANT: Lowman, Henry
: TITLE OF INVENTION: PROTEIN VARIANTS
: FILE REFERENCE: P171281-1
: CURRENT APPLICATION NUMBER: US/10/028,410
: CURRENT FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: US/03/477,424
: PRIOR FILING DATE: 2000-01-15
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 3
: LENGTH: 51
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-028-410-3

Query Match 94.74: Score 278.5: DB 10: Length 124:
Best Local Similarity 94.34: Pred. No. 1,4e-26:
Matches 51: Conservative 0: Mismatches 0: Indels 1: Gaps 1:

QY : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 51

RESULT 6
US-09-894-711-18
: Sequence 18, Application US/09894711
: Patent No. US20020137144A1
: GENERAL INFORMATION:
: APPLICANT: Kjeldsen, Thomas Borglum
: APPLICANT: Ludvigsen, Svend
: TITLE OF INVENTION: Method for making insulin precursors and
: TITLE OF INVENTION: insulin precursor analogues having improved formulation
: TITLE OF INVENTION: yield in yeast
: FILE REFERENCE: 6149,400-US
: CURRENT APPLICATION NUMBER: US/09/894,711
: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: PA 2000 00443
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: PA 1999 01869
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: 63/211,081
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: 63/181,450
: PRIOR FILING DATE: 2000-02-10
: PRIOR APPLICATION NUMBER: 63/740,359
: PRIOR FILING DATE: 2000-12-19
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 18
: LENGTH: 124
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic
US-09-894-711-18

Query Match 93.74: Score 275.5: DB 10: Length 124:
Best Local Similarity 94.34: Pred. No. 1,4e-26:
Matches 50: Conservative 1: Mismatches 1: Indels 1: Gaps 1:

QY 1 FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB 72 FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 124

RESULT 7
US-09-861-687-19
: Sequence 19, Application US/C9861687
: Publication No. US20020193292A1
: GENERAL INFORMATION:
: APPLICANT: Markussen, Jan
: APPLICANT: Jonassen, Ib
: APPLICANT: Brandt, Jakob
: APPLICANT: Kurtzhals, Peter
: APPLICANT: Hansen, Hertz Per
: APPLICANT: Kaarsholm, Niels Christian
: TITLE OF INVENTION: INSULIN DERIVATIVES
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NO. US20020193292A1o No. US20020193292A1disk of No. US20
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/09/861,687
: APPLICATION NUMBER: US/09/861,687
: FILING DATE: 21-May-2001
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 68/932,082
: FILING DATE: 16-DEC-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 4341,204-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 138 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-861-687-19

Query Match 93.74: Score 275.5: DB 10: Length 138:
Best Local Similarity 94.34: Pred. No. 1,4e-26:
Matches 50: Conservative 1: Mismatches 1: Indels 1: Gaps 1:

QY 1 FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB 86 FVNQHLGGSHLVKALVCGGEGFFYTPKTRGIVEQCCTSCISLYOLENYCN 138

RESULT 8
US-10-056-009A-3
: Sequence 3, Application US/10066009A
: Publication No. US20020165155A1
: GENERAL INFORMATION:
: APPLICANT: Schaffer, Michelle
: APPLICANT: Jitsos, Mark
: APPLICANT: Vajdos, Felix
: TITLE OF INVENTION: CRYSTALLIZATION OF IGF-1
: FILE REFERENCE: P1669R1

```
/ CURRENT APPLICATION NUMBER: US/06/066,89A
/ 
/ CURRENT FILING DATE: 2002-06-24
/ PRIOR APPLICATION NUMBER: US 60/287,072
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/267,977
/ PRIOR FILING DATE: 2001-02-00
/ NUMBER OF SEQ ID NOS: 5
/ SEQ ID NO 1
/ LENGTH: 50
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-066-C09A.1

Query Match      92.9%   Score 273    DP 17   Length 50
Best Local Similarity 96.2%   Pred. No. 9,4e+27
Matches 50: Conservative C; Mismatches G; Indels 2; Gaps 1;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYYPKTKGLIVQECCSTISCSLYLENVEN 52
|||||.....|||.|||||..I.....|.....|.....|
Db 1 FVNHLGGSHLVEALYLVCGERGFFYYPK--GIIVEQCCTISCSLYLENVEN 50
|||||.....|||.|||||..I.....|.....|.....|

RESULT 9
US-09-878-380-1
Sequence 1, Application US/09878380
Patent No. US20020160435A1
GENERAL INFORMATION:
APPLICANT: Fulirebio Inc.
APPLICANT: KITAHIMA, Sachiko
APPLICANT: KURANO, Yoshihiro
APPLICANT: NAKATSUBO, Kaoru
APPLICANT: NISHIZONO, Isao
TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Thereof
FILE REFERENCE: 0760-029P
CURRENT APPLICATION NUMBER: US/09/878,380
CURRENT FILING DATE: 2001-06-12
PRIOR APPLICATION NUMBER: JP 2000-174691
PRIOR FILING DATE: 2000-06-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-09-878-380-1

Query Match      90.8%   Score 267    DP 17   Length 86
Best Local Similarity 90.5%   Pred. No. 9,4e+26
Matches 52: Conservative C; Mismatches G; Indels 3; Gaps 1;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYYPKT----- ..... 40
|||||.....|||.|||||.....|.....|.....|
Db 1 FVNHLCGSHLVEALYLVCGERGFFYYPKTGKRAEHLGVQLVVLEGGSPHNSIQFLAD 60
|||||.....|||.|||||.....|.....|.....|
QY 31 ----RGIVEQCCTISCSLYLENVEN 52
||.....|.....|.....|.....|
Db 61 SLOKGRIEQCCTISCSLYLENVEN 55
||.....|.....|.....|.....|

RESULT 10
US-09-858-935B-4
Sequence 4, Application US/C9358935B
Publication No. US20030069177A1
GENERAL INFORMATION:
APPLICANT: Dunaquie, Yves
APPLICANT: Filvarciff, Ellen
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P'79431
CURRENT APPLICATION NUMBER: US/09/858,935B
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/248,905
PRIOR FILING DATE: 2000-11-13
```

```
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA: US/10/054,873
? APPLICATION NUMBER: US/10/054,873
? FILING DATE: 22-Jan-2002
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/CN98/00052
? FILING DATE: 31-Mar-1998
? APPLICATION NUMBER: US 09/423,100
? FILING DATE: 11-DEC-2003
? ATTORNEY/AGENT INFORMATION:
? NAME: Mycroft, Frank J
? REGISTRATION NUMBER: 45,646
? REFERENCE/DOCKET NUMBER: 020167-000130/5
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 86 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4

Query Match 90.8% Score 267; DB 14; Length 86;
Best Local Similarity 60.5%; Pred. No. 1.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 86
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 86
QY 31 ----RGIVECCCTSCISLYOLENYCN 52
DB 61 SLQKRGIVECCCTSCISLYOLENYCN 86

RESULT 13
US-09-947-563-4
? Sequence 4, Application US/09/47563
? Patent No. US20020156234A1
? GENERAL INFORMATION:
? APPLICANT: Subroder, Franz-Josef
? Keller, Reinhold
? TITLE OF INVENTION: Improved process for obtaining
? insulin precursors having correctly bonded cystine bridges
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS
? ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
? Dunner
? STREET: 1400 I Street, N.W.
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/947,563
? FILING DATE: 07-Sep-2001
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA
? APPLICATION NUMBER: 09/134,636
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Leslie McDonnell
? REGISTRATION NUMBER: 34,872
? REFERENCE/DOCKET NUMBER: 02481.1500-00000
```

```
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 408-4000
? TELEFAX: (202) 408-4400
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 96 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? ORIGINAL SOURCE:
? ORGANISM: Escherichia coli
? FEATURE:
? NAME/KEY: Protein
? LOCATION: 1..96
? SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

Query Match 90.8% Score 267; DB 10; Length 96;
Best Local Similarity 60.5%; Pred. No. 1.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 30
DB 21 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 70
QY 31 ----RGIVECCCTSCISLYOLENYCN 52
DB 71 SLQKRGIVECCCTSCISLYOLENYCN 96

RESULT 14
US-09-205-658-125
? Sequence 125, Application US/09205658
? Patent No. US20010029617A:
? GENERAL INFORMATION:
? APPLICANT: Ruvkun, Gary
? APPLICANT: Ogg, Scott
? TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
? TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
? FILE REFERENCE: 00786/351004
? CURRENT APPLICATION NUMBER: US/09/205,658
? CURRENT FILING DATE: 1998-12-03
? EARLIER APPLICATION NUMBER: 08/857,876
? EARLIER FILING DATE: 1997-05-15
? EARLIER APPLICATION NUMBER: 08/886,534
? EARLIER FILING DATE: 1997-07-07
? EARLIER APPLICATION NUMBER: US98/10080
? EARLIER FILING DATE: 1998-05-15
? NUMBER OF SEQ ID NOS: 328
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 125
? LENGTH: 110
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-205-658-125

Query Match 90.8% Score 267; DB 9; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.2e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 30
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 84
QY 31 ----RGIVECCCTSCISLYOLENYCN 52
DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 15
US-09-615-229-3
? Sequence 3, Application US/09815229
```

```

: Patent No. US2002005862A1
: GENERAL INFORMATION:
: APPLICANT: Filvaroff, Elion B.
: APPLICANT: Okuno, Franklin W.
: TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARCINOGENESIS IN SKELETON
: FILE REFERENCE: P1786R1US
: CURRENT APPLICATION NUMBER: US 09/815,229
: CURRENT FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: US 60/52,103
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 17
: SEQ ID NO 3
: LENGTH: 116
: TYPE: PRT
: ORGAN: SM: Homo sapiens
US-09-815-229-3

```

```

Query Match      90.88; Score 257; Hs 9; Length 116;
Best Local Similarity 60.58; Pred. No. 1,20-25;
Matches 52; Conservative 0; Mismatches 9; Gaps 11
QY      1  FVNQLCGSHLVEALYLVGGRGFFYTKT----- -- 30
Db      1  FVNQLCGSHLVEALYLVGGRGFFYTKT----- -- 30
QY      25 FVNQLCGSHLVEALYLVGGRGFFYTKT----- -- 30
Db      25 FVNQLCGSHLVEALYLVGGRGFFYTKT----- -- 30
QY      31 ----RGIVECCCTSTCSLFOLENYCN 52
Db      31 ----RGIVECCCTSTCSLFOLENYCN 52
QY      35 SLCKRGIVECCCTSTCSLFOLENYCN 110
Db      35 SLCKRGIVECCCTSTCSLFOLENYCN 110

```

Search completed: September 15, 2003, 12:23:36
Job time : 22.8065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Inc.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 : Search time 8.67348 seconds
(without alignments)
583.284 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 FVNHQGLSHLVLEALVCG.....IVEQCCTICSLYLENYCN 52

Scoring table: BLOSUM62
Gapof 10.0 , Gapext 0.5

Searched: 283308 seqs, 96162682 residues
Total number of hits satisfying chosen parameters: 283908

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database : PTR_76:★
1: p1r1:★
2: p1r2:★
3: p1r3:★
4: p1r4:★

Pred. No. is the number of results predicted by change to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273.5	93.0	51	INMSP	insulin - sperm whale
2	273.5	93.0	51	INMHP	insulin - finback
3	273.5	93.0	51	INM	insulin - elephant
4	273.5	92.9	96	PC7582	epidermal growth f
5	271.5	92.3	51	ENY	insulin - hamster
6	268.5	91.3	51	INMSSP	insulin - Egyptian
7	267.5	91.0	51	A59151	insulin precursor
8	267	90.8	120	IPHC	insulin precursor
9	267	90.8	110	B42179	insulin precursor
10	267	90.8	110	A42179	insulin precursor
11	267	90.8	110	CQ1178	insulin precursor
12	263.5	89.6	51	INMHL	insulin - sei whale
13	263.5	89.6	51	INMHL	insulin - goat
14	263.5	89.6	51	INMKA	insulin - Arabian
15	263	89.5	84	IPG	insulin precursor
16	263	89.5	110	INM	insulin precursor
17	262.5	89.3	51	INM	insulin - rat
18	262	89.1	110	IPG	insulin precursor
19	261.5	88.9	51	INMKS	insulin - common s
20	260	88.4	110	I48166	insulin precursor
21	258.5	87.9	135	IPRO	insulin precursor
22	256.5	87.2	51	JQ0362	insulin - North Am
23	252.5	85.9	77	INSH	insulin precursor
24	252	85.7	86	IPHC	insulin precursor
25	251.5	85.5	51	INM	insulin - Chicchi
26	251	85.4	128	A36883	insulin precursor
27	250	85.0	108	INM51	insulin precursor
28	249	84.7	110	IPRT1	insulin precursor
29	248.5	84.5	51	INM	insulin - deorse

30	248	84.4	110	1	IPRT2	insulin 2 precursor
31	248	84.4	110	1	INM52	insulin 2 precursor
32	246	83.7	52	2	S44669	insulin 11 - North
33	246	83.7	52	2	S44670	insulin 12 - North
34	244.5	83.2	51	1	INPQ	insulin - crested
35	244.5	83.2	51	1	INTK	insulin - turkey
36	244.5	83.2	51	1	INOS	insulin - ostrich
37	244.5	83.2	51	1	A61129	insulin - black-be
38	244.5	83.2	51	2	A60414	insulin - slider t
39	238.5	81.1	103	2	IS1221	insulin precursor
40	238	81.0	52	2	S61361	insulin - Amphiuma
41	235.5	80.1	51	2	S63590	insulin - duckbill
42	234.5	79.8	107	1	IPCH	insulin precursor
43	233.5	79.4	81	1	IPDK	insulin precursor
44	231.5	78.7	51	1	INAO	insulin - American
45	231	78.6	52	1	INMKA	insulin - alligato

ALIGNMENTS

RESULT 1

INMHP
insulin - sperm whale
C.Species: Physeter catodon (sperm whale)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C.Accession: A93142: A90082
R.Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 161, 1466-1469, 1958
A.Title: Structure of sperm- and sei-whale insulins and their breakdown by whale I
A.Reference number: A93142
A.Accession: A93142
A.Molecule type: protein
A.Residues: 1-30:31-51<ISH>
R.Barris, C.I.; Sanger, F.; Naughton, M.A.
Arch. Biochem. Biophys. 85, 427-428, 1956
A.Title: Species differences in insulin.
A.Reference number: A90082
A.Accession: A90082
A.Molecule type: protein
A.Residues: 1-30:31-51<HAR>

Query Match 93.0% Score 273.5; DB 1; Length 51;
Best local similarity 96.2%; Pred. No. 1.5e-24;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

RESULT 2

INMHP
insulin - finback whale (tentative sequence)
C.Species: Balenoptera physalus (finback whale, common rorqual)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C.Accession: A91918
R.Hama, H.; Titani, K.; Sakaki, S.; Narita, K.
C. Biochem. 56, 265-293, 1964
A.Title: The amino acid sequence in fin-whale insulin.
A.Reference number: A91918
A.Accession: A91918
A.Molecule type: protein
A.Residues: 1-30:31-51<HAM>
C.Superfamily: insulin
C.Keywords: hormone; pancreas
F.1-30/Domain: insulin chain B #status experimental <BCH>

A:Accession: B42179
A:Molecule type: cDNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:q38251; PID:q38251
A:Note: Sequence extracted from NCB1 backbone (NID:q38251; NID:q38251)
A:Reference: J. Biol. Chem. 247, 4865-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and bat proinsulin
A:Reference number: A92111; MUI:172258016; PMID:462369
A:Accession: A05232
A:Molecule type: protein
A:Residues: 57-87 <PPT>
A:Genetics:
A:Introns: 63/
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain A #status predicted <ACH>
F:55-87/Domain: insulin chain B #status predicted <ACH>
F:57-87/Domain: connecting peptide #status experimental <CHP>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:91-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 90.88; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 45
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRRAEDQVGVQLGGGPGAGSLQPLALEG 44

QY 31 ----RGIVEQCCTSCISLYQLENYCN 52
DB 85 SLQKRGIVEQCCTSCISLYQLENYCN 110

RESULT 10
A42179
insulin precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: A42179; S22058
R:Seino, S.; Reilly, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate of evolution in the human lineage
A:Reference number: A42179; MUI:92215953; PMID:11560757
A:Accession: A42179
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:q38251; PID:q38251; PMID:462369
A:Note: Sequence extracted from NCB1 backbone (NID:q38251; NID:q38251)
A:Genetics:
A:Introns: 63/
A:Superfamily: insulin

Query Match 90.88; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 45
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRRAEDQVGVQLGGGPGAGSLQPLALEG 44

QY 31 ----RGIVEQCCTSCISLYQLENYCN 52
DB 85 SLQKRGIVEQCCTSCISLYQLENYCN 110

RESULT 11
J00178
insulin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C:Accession: J00178
A:Molecule type: cDNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:q38251; PID:q38251
A:Reference: J. Biol. Chem. 247, 4865-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and bat proinsulin
A:Reference number: A92111; MUI:172258016; PMID:462369
A:Accession: A05232
A:Molecule type: protein
A:Residues: 57-87 <PPT>
A:Genetics:
A:Introns: 63/
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain A #status predicted <ACH>
F:55-87/Domain: insulin chain B #status predicted <ACH>
F:57-87/Domain: connecting peptide #status experimental <CHP>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:91-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 90.88; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 45
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRRAEDQVGVQLGGGPGAGSLQPLALEG 44

QY 31 ----RGIVEQCCTSCISLYQLENYCN 52
DB 85 SLQKRGIVEQCCTSCISLYQLENYCN 110

RESULT 12
J00178
insulin precursor - sei whale
C:Species: Balaeonotera borealis (sei whale)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C:Accession: A01582
R:Stihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 281, 1468-1469, 1978
A:Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pe
A:Reference number: A03142
A:Accession: A01582
A:Molecule type: protein
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:q38251; PID:q38251; PMID:462369
A:Note: Sequence extracted from NCB1 backbone (NID:q38251; NID:q38251)
A:Genetics:
A:Introns: 63/
A:Superfamily: insulin

Query Match 89.6%; Score 263.5; DB 1; Length 52;
Best Local Similarity 92.3%; Pred. No. 2.1e-23;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 52
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKA-GIVEQCCTSCISLYQLENYCN 51

RESULT 13
J00178
insulin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Jul-1999
C:Accession: A01586
R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUI:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-110 <SEI>

Search completed: September 15, 2003, 12:04:20
Job time : 9.57346 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw mode

Run on: September 15, 2003, 11:54:00 : Search time 4.81588 seconds
(without alignment)
524.033 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 234
Sequence: 1 FVQHICGSHLVEALYLVCG.....IVEQCCTSIGSLYLENYCN 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127853 seqs, 47025705 residues

Total number of hits satisfying chosen parameters: 127853

Minimum DB seq length: 0

Maximum DB seq length: 2300000000

Post-processing: Minimum Match 1%

Listing first 15 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	273.5	93.0	51	1	INS_HALPH	Human insulin
2	271.5	91.0	51	1	INS_ELEMA	Elephant insulin
3	268.5	91.3	51	1	INS_ACOCA	Porcine insulin
4	267.0	90.8	110	1	INS_GERAP	Guinea pig insulin
5	267.0	90.8	110	1	INS_HOVAN	Human insulin
6	267.0	90.8	110	1	INS_MACFA	Macaca insulin
7	267.0	90.8	110	1	INS_PANTR	Human insulin
8	263.5	89.6	51	1	INS_RALBO	Human insulin
9	263.5	89.6	51	1	INS_GAMOR	Human insulin
10	263.5	89.6	51	1	INS_GAPHI	Human insulin
11	263.5	89.6	51	1	INS_PEG	Human insulin
12	263.5	89.5	108	1	INS_PABIT	Human insulin
13	263.5	89.5	110	1	INS_SPTIR	Human insulin
14	262.5	89.3	51	1	INS_FELCA	Human insulin
15	262.5	89.1	110	1	INS_CANEA	Human insulin
16	260.5	88.4	110	1	INS_CRILC	Human insulin
17	258.5	87.9	105	1	INS_BOVIN	Human insulin
18	257.5	87.4	110	1	INS_PSAOH	Human insulin
19	256.5	87.2	51	1	INS_DIDMA	Human insulin
20	254.5	86.6	105	1	INS_SHEEP	Human insulin
21	252.5	85.7	86	1	INS_FORSE	Human insulin
22	251.5	85.5	51	1	INS_CHIBR	Human insulin
23	251.5	85.4	108	1	INS_AOVPR	Human insulin
24	250.5	85.0	108	1	INS_MOUSE	Human insulin
25	249.5	84.7	110	1	INS_I_RAT	Human insulin
26	248.5	84.5	51	1	INS_ANGAN	Human insulin
27	248.5	84.4	110	1	INS2_MOUSE	Human insulin
28	248.5	84.4	110	1	INS2_RAT	Human insulin
29	246.5	83.7	52	1	INS_ACIGU	Human insulin
30	244.5	83.2	51	1	INS_HYSCR	Human insulin
31	244.5	83.2	51	1	INS_TRASC	Human insulin
32	238.5	81.1	103	1	INS_SELRF	Human insulin
33	235.5	80.1	51	1	INS_ORNAN	Human insulin

33 234.5 79.8 107 1 INS_CHICK
35 233.5 78.4 81 1 INS_ANAPL
36 231.5 78.7 51 1 INS_ALAM:
37 231.5 78.6 52 1 INS_LEPSP
38 228.5 77.7 51 1 INS_ZAODH
39 227.5 77.4 51 1 INS_CROAT
40 226.5 77.0 106 1 INS1_XENLA
41 226.5 77.0 106 1 INS2_XENLA
42 222.5 75.5 57 1 INS_PEMA
43 216.5 73.5 51 1 INS_PLAPE
44 214.5 72.8 52 1 INS_AMICA
45 211.5 71.8 50 1 INS_ONCGO

ALIGNMENTS

RESULT 1
TE INS_HALPH STANDARD; PRT: 51 AA.
AC P01312;
DE 21-JUN-1986 (Rel. 01, Created)
DC 21-JUN-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
ID Insulin.
IN INS.
OS Balaenoptera physalus (Finback whale) (Common rorqual), and
US Physeter catodon (Sperm whale) (Physeter macrocephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
CX NCBI_FaxID:9770, 9755;
FN 1;
PP PARTIAL SEQUENCE.
RC SPECIES=B.physalus;
RA Hama H., Titani K., Sakaki S., Narita K.;
RT "The amino acid sequence in fin-whale insulin.";
RL J. Biochem. 56:285-293(1964).
RN 1;
RO SEQUENCE.
RA Ishihara Y., Saito J., Ito Y., Fujino M.;
RT "Structure of sperm- and sei-whale insulins and their breakdown by
KL whale pepsin.";
RL Nature 181:448-449(1958).
RN 1;
PP SEQUENCE.
RC SPECIES=P.catodon;
RA Harris J., Sanger F., Naughton M.A.;
RT "Species differences in insulin.";
RL Arch. Biochem. Biophys. 65:427-439(1956).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A91918; INMHP.
DR PIR: A93142; INMHP.
DR RSP: P01317; IAPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SMC0078; IIGF; 1
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NGL_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA: 5766 MW: 5007B514691A7CDD CRC64:

```

NCBI_TAXID=10968;
[1]
COMPOSITION.
MEDLINE=7258454; PubMed=5028210;
Bueneli H.F., Humbel R.E.;
"Isolation and partial structural analysis of insulin from mouse (Mus musculus) and spiny mouse (Acomys kuhnyi).";
Poppe-Sejler's Z. Physiol. Chem. 353:444-450(1972);
-!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
-!- CYCLIN: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
PIR: A01591; INKSSP.
PIR: P01308; ITRY.
Zlotar'ko; IPR004825; Ins/IGF/relax.
SWARC: SM00078; IIGF; J.
PROSITE; PS00262; INSULIN; 1.
Insulin family; Hormone; Glucose metabolist.
CHAIN 1 30 INSULIN B CHAIN.
FT NON-CONS 30 31
FT CHAIN 31 51 FT CHAIN
FT DISULFID 7 37
FT DISULFID 19 50
FT DISULFID 36 41
FT DISULFID 36 41
SEQUENCE 51 AA: 5768 MW: 9928086629047030 CRC64;
Query Match 91.3%; Score 268.5; DB 1; Length 51;
Best Local Similarity 92.3%; Prid. No. 28-26;
Matches 48; Conservative 1; Mismatches 0; Indels 1; Gaps
CY 1 FVNHQLGSHLVEALYLVCGERGFYTPKTRGIVGQVCTSCSYLYOLENYCN 52
DB 1 FVHQLGSHLVEALYLVCGERGFYTPKS-GIVDQCTSCSYLYOLENYCN 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 4
INS_CERAE
AC P30407; P01309; STANDARD; PRT; 110 AA.
DT 01-APR-1993 (Rel. 25, Created)
ET 01-APR-1993 (Rel. 25, Last sequence update)
ST 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
CC Cercopithecus aethiops (Green monkey) (Grivet).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecinae; Cercopithecus.
OX NCBI_TAXID=9534;
[1]
SEQUENCE FROM N.A.
EX MEDLINE=92219953; PubMed=5560757;
SA Seino S., Bell G.I., Li W.;
FT "Sequences of primate insulin genes support the hypothesis of a slower rate of molecular evolution in humans and apes than in monkeys.";
ST Mol. Biol. Evol. 9:193-203(1992).
[2]
SEQUENCE OF 57-87.
EX MEDLINE=72258016; PubMed=4626369;
SA Peterson S.D., Nehrlich S., Oyer P.E., Steiner D.F.;
FT "Determination of the amino acid sequence of the monkey, sheep, and dog proinsulin C-peptides by a semi-micro Edman degradation procedure.";
SA J. Biol. Chem. 247:4866-4871(1972).
-!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.

```


RP SYNTHESIS OF 57-57.
RX MEDLINE=75540007; PubMed=4923304;
RA Naichani V.K.;
RT "Studies on polypeptides. IV. The synthesis of C peptide of human
RT proinsulin";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:155-162(1973).
RN [13]
RP SYNTHESIS OF 45-69 AND 70-73.
RX MEDLINE=73161263; PubMed=4598559;
RA Geiser R., Volk A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). 3. Synthesis of the sequences 44-17 and 6-14
RT of human proinsulin C peptides";
RL Chem. Ber. 106:199-205(1973).
RN [14]
RP SYNTHESIS OF 84-87
RX MEDLINE=73161261; PubMed=4598553;
RA Geiger R., Jaeger G., Keonig W., Treuth G.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). I. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide";
RL Chem. Ber. 106:188-192(1973).
RN [15]
RP VARIANT LOS ANGELES SER-46.
RX MEDLINE=84316053; PubMed=6212455;
RA Haceda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding (SerB34, Insulin)";
RL Proc. Natl. Acad. Sci. U.S.A. 50:6366-6370(1983).
RN [16]
RP VARIANTS LOS ANGELES SER-45 AND CHICAGO LEU-49.
RX MEDLINE=84170233; PubMed=6424111;
RA Shoelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.F.;
RT "Identification of a mutant human insulin predicted to contain a
RT serine-for-phenylalanine substitution";
RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RN [17]
RP VARIANT PROVIDENCE ASP-34.
RX MEDLINE=87175640; PubMed=3470784;
RA Chan S.J., Seino S., Gruppiso P.A., Schwartz R., Steiner D.F.;
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
RN [18]
RP VARIANT WAKAYAMA LEU-52.
RX MEDLINE=87058122; PubMed=3337011;
RA Sakura H., Tamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A5 (Val¹ Glu) isolated from the pancreas";
RL J. Clin. Invest. 78:1666-1672(1986).
RN [19]
RP VARIANT HIS-89.
RX MEDLINE=90317021; PubMed=2196279;
RA Harbetti F., Raben N., Kadowaki T., Cama A., Arcill E., Gaby K.H.;
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 89 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction";
RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
RN [20]
RP VARIANT HIS-89.
RX MEDLINE=85261996; PubMed=4019786;
RA Shibasaki Y., Kawakami T., Kanazawa Y., Akazawa Y., Takazu F.;
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia";
RL J. Clin. Invest. 76:378-380(1985).
RN [21]
RP VARIANT KYOTO LEU-83.
RX MEDLINE=92291307; PubMed=1401997;
RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Iwata H., Seino Y.;
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto)";
RL J. Clin. Invest. 89:1902-1907(1992).
RN [22]
RP STRUCTURE BY NMR.
RX MEDLINE=91104966; PubMed=2271664;
RA Hua O.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RT structure";
RL Biochemistry 29:10545-10555(1996).
RN [23]
RP STRUCTURE BY NMR.
RX MEDLINE=91242467; PubMed=2036420;
RA Hua O.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition";
RL Biochemistry 30:5505-5515(1991).
RN [24]
RP STRUCTURE BY NMR.
RX MEDLINE=91265527; PubMed=1646635;
RA Hua O.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Des-(B26-530)-insulin: sequence-
RT specific resonance assignments and effects of solvent composition";
RL Biochim. Biophys. Acta 1078:101-110(1991).
RN [25]
Query Match 90.8%; Score 267; DB 1; Length 110;
Eust Local Similarity 60.5%; Pred. No. 6.3e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
QY 1 FVNRGSGSLVAVLVGGERGFYTPKT----- 30
DB 25 FVNRGSGSLVAVLVGGERGFYTPKTREAEVLGVGVGGGAGSLOPLALEG 84
QY 31 ----RGIVGCGCTSCSLYQLENYCN 52
DB 85 SLOKRGIVGCGCTSCSLYQLENYCN 110
RESULT 6
INS_MACFA
ID INS_MACFA STANDARD; PRT: 110 AA.
AC P30406; P01309;
DT 21-JUL-1986 (Rel. 01, Created)
DT 23-AUG-1987 (Rel. 05, Last sequence update)
LT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS
CS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecinae; Macaca.
CC NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83083474; PubMed=6184262;
RA Wetkham W., Gronenberg E., Heineweb M., Wengenmayer F.,
RA Winnacker E.-L.;
RT "The nucleotide sequence of cDNA coding for proinsulin from the
RT primate Macaca fascicularis";
RL Gene 19:179-183(1982).
RN [2]
RP FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
RP INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
RP FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
RP CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
RP SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
RP DISULFIDE BONDS.
RP CELLULAR LOCATION: Secreted.
RP SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/about-us/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: J00336; AAA36849.1; -
 DR PIR: Q0178; JQ0178.
 DR HSSP: P01408; IAT0.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS: PR00277; INSULIN.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 91 96 INTERCHAIN.
 FT DISULFID 43 109 INTERCHAIN.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA: 1199; MW: 836633AP04A20F9 C6C64;

Query Match 90.8% Score 267; DR 1; Length 110;
 Best Local Similarity 60.5% Pred. No. 6.3e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNOHCGSHLVEALYVGCGEGFYTPKT ----- 30
 DB 25 FVNOHCGSHLVEALYVGCGEGFYTPKTREAEALQVGVFVGCGGPGAGSLQPLAEG 84
 QY 31 ----RGIVEGCGTSGISLYOLENYCN 52
 DB 85 SLGKPGIVEGCGTSGISLYOLENYCN 110

RESULT 7
 INS_PANTR
 ID INS_PANTR STANDARD; PRT; 110 AA.
 AC P30410;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Insulin precursor.
 GN INS.
 OS Pan troglodytes (Chimpanzee).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 CX NCBI_TaxID=9598;
 RN [1]
 RP MEDLINE:92219953; PubMed:1560757;
 RA Seino S., Bell G.I., Li W.
 RT "Sequences of primate insulin genes support the hypothesis of a
 RT slower rate of molecular evolution in humans and apes than in
 RT monkeys."
 RL Mol. Biol. Evol. 9:193-203(1992).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SURUNIT: HETEROIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/about-us/>

CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X61060; CAA43403.1; -
 DR PIR: A42179; A42179.
 DR DB: IEP; 29-MAR-00.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS: PR00277; INSULIN.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 91 96 INTERCHAIN.
 FT DISULFID 43 109 INTERCHAIN.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA: 12025 MW: 41E8DF79837CEEF5 C6C64;

Query Match 90.8% Score 267; DR 1; Length 110;
 Best Local Similarity 60.5% Pred. No. 6.3e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNOHCGSHLVEALYVGCGEGFYTPKT ----- 30
 DB 25 FVNOHCGSHLVEALYVGCGEGFYTPKTREAEALQVGVFVGCGGPGAGSLQPLAEG 84
 QY 31 ----RGIVEGCGTSGISLYOLENYCN 52
 DB 85 SLGKPGIVEGCGTSGISLYOLENYCN 110

RESULT 8
 INS_BALBO
 ID INS_BALBO STANDARD; PRT; 51 AA.
 AC P31334;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Insulin.
 GN INS.
 OS Balanoptera borealis (Sei whale).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
 CC Balanopteridae; Balanoptera.
 CX NCBI_TaxID=3768;
 RN [1]
 RP SEQUENCE.
 RA Ishihara Y., Saito T., Ito Y., Fujico M.;
 RC "Structure of sperm- and sei-whale insulins and their breakdown by
 RC whale pepsin."
 RL Nature 181:1468-1479(1958).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SURUNIT: HETEROIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC P-R: A01582; INWHIS.
 DR HSSP: P01317; IAPH.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism.
 FT CHAIN 1 30 INSULIN B CHAIN.
 FT NON-CONS 30 31
 FT CHAIN 31 51 INSULIN A CHAIN.
 FT DISULFID 7 37 INTERCHAIN.
 FT DISULFID 19 50 INTERCHAIN.
 FT DISULFID 36 41

[illegible]

RP REVISION TO 59.
 RA Change R.E.;
 RL Submitted (JUL-1970) to the PIR data bank.
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RA Blundell I.L., Dodson G.G., Hodgkin D., Moras D.J.
 RL "Insulin. The structure of the crystal and its reflection in
 chemistry and biology."
 RL Adv. Protein Chem. 26:279-402(1972).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RA Isaacs N.W., Adarwa R.C.;
 RL "Experience with fast Fourier least squares in the refinement of the
 crystal structure of rhombohedral 2-zinc insulin at 1.5-A
 resolution."
 RL Acta Crystallogr. A 34:782-791(1978).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RA Baker E.N., Blundell I.L., Cutfield J.F., Dodson G.G.,
 RA Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
 RA Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
 RL "The structure of 220 pig insulin crystals at 1.5-A resolution."
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-436(1978).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RA MEDLINE-92126280; PubMed-1772543;
 RA Baltschmidt P., Hansen F.B., Dodson G., Korber F.;
 RL "Structure of porcine insulin cocrystallized with elipatine Z."
 RL Acta Crystallogr. B 47:975-986(1991).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY.
 RA MEDLINE-9122450; PubMed-2025410;
 RA Badger G., Harris M.R., Reynolds C.D., Evans A.C., Dodson G.G.;
 RA Dodson G.G., North A.C.F.;
 RL "Structure of the pig insulin dimer in the cubic crystal."
 RL Acta Crystallogr. B 47:127-136(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
 RA Zhao J.-S., Wan Z.-L., Chang W.-K., Liang D.-Q.;
 RL "Structure of monomeric porcine Des¹⁻²² B2 desptenapeptide (326-540)
 insulin at 1.55-A resolution."
 RL Acta Crystallogr. B 53:507-512(1997).
 CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC -1- DATABASE: NAME-Protein Spotlight;
 CC NOTE-Issue 9 of April 2003;
 CC WWW="http://www.expasy.org/spotlight/articles/spotlight.html".

DR PCB: IMED: 29-JAN-96.
 DR DB: 3MCH: 29-JAN-96.
 DR EUB: IDEI: 16-JAN-97.
 DR PCB: ISDB: 01-APR-98.
 DR DB: IWAY: 28-FEB-97.
 DR FDB: IZEL: 16-FEB-99.
 DR DB: IZNI: 28-JAN-98.
 DR PCB: IZNI: 28-JAN-98.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL; 24
 FT CHAIN; 25 INSULIN B CHAIN.
 FT PROPEP; 57 85 C-PEPTIDE.
 FT CHAIN; 86 106 INSULIN A CHAIN.
 FT DISULFID; 31 94 INTERCHAIN.
 FT DISULFID; 43 107 INTERCHAIN.
 FT DISULFID; 93 98
 FT HELIX; 26 46
 FT STRAND; 48 48
 FT STRAND; 89 94
 FT HELIX; 100 106
 FT STRAND; 107 107
 SC SEQUENCE 108 AA; 11671 MW; CB44918429658EBC CRC64;
 Query Match 89.5%; Score 263; DB 1; Length 108;
 PestLocal: Similarity 60.7%; Pred. No. 1.9e-25;
 Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1;
 QY 1 FVNHQHGSHLVEALYVCGERGFFYTPKTI----- 30
 DB 25 FVNHQHGSHLVEALYVCGERGFFYTPKARREAPNGAVELGGISGQAALAEQPP 84
 QY 31 ---KGIVEQCCTSGSLYQLENYCN 52
 DB 25 QKRGIVEQCCTSGSLYQLENYCN 108
 RESULT 12
 INS_RABIT STANDARD PRI: 110 AA.
 AC P01311;
 DE 21-JUL-1986 (rel. 01, Created)
 DE 01-FEB-1996 (rel. 33, Last sequence update)
 DE 01-OCT-1996 (rel. 34, Last annotation update)
 DE Insulin precursor.
 GN INS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 ON NCBI_TaxID=9946;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-New Zealand white; Tissue-Pancreas;
 RX MEDLINE-94179230; PubMed-8132571;
 RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
 RA Menon S.K., Zahm D.S.;
 RL "Insulin gene expression and insulin synthesis in mammalian neuronal
 cells."
 RL J. Biol. Chem. 269:8445-8454(1994).
 RN [2]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE-66160119; PubMed-5949593;
 RA Smith L.F.;
 RL "Species variation in the amino acid sequence of insulin."
 RL Am. J. Med. 40:662-666(1966).
 RN [3]
 RP SEQUENCE OF 56-110 FROM N.A.
 RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
 RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT

```

CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch.)
CC -----
DR EMBL: U33610; AAA15033.1;
DR EMBL: M61253; AAA17540.1;
DR PIR: A53438; INPH.
DR HSSP: P01308; ITHM.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin.
DR SMART: SM00078; IIGF.
DR PROSITE: PS00262; INSULIN.
DR Insulin family: Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24 INSULIN B CHAIN.
FT CHAIN 25 54 C PEPTIDE.
FT PROPEP 57 87
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 33 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 85 E -> Y (IN REF. 3).
SQ SSSEQUENCE 110 AA: 11818 MW: 826297585D77EA8 CRC64:

Query Match 89.5%; Score 263; DB 1; Length 110;
Best Local Similarity 59.3%; Pred. No. 1.9e-25;
Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNHLCGSHVEALYLVGGERGFYTPK----- 30
Db 25 FVNHLCGSHVEALYLVGGERGFYTPKSRREVERELVGGAEIGGPGAGLQPSALEL 84
QY 31 ----RGIVGCGCTSGSLYOLENYCN 52
Db 85 ALOKRGIVEGCTSGSLYOLENYCN 110

RESULT 13
INS_SPETR
ID INS_SPETR STANDARD: PRT: 110 AA.
AC Q91X13;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
DE INS.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spermophilus.
OC NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Pancreas;
RA Tredrea M.M., Buck M.J., Ghanalyogi J., Squire P.D., Andrews K.L.;
RT "Regulation of IGF4 expression in a hibernating mammal."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

```

```

CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch.)
CC -----
DR EMBL: AY338604; AAK72558.1;
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin.
DR PROSITE: PS00277; INSULIN.
DR SMART: SM00078; IIGF.
DR PROSITE: PS00262; INSULIN.
DR Insulin family: Hormone; Glucose metabolism; Signal.
FT SIGNAL 24 BY SIMILARITY.
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 94 INTERCHAIN (BY SIMILARITY).
FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY).
FT DISULFID 95 100 BY SIMILARITY.
SQ SEQUENCE 110 AA: 12004 MW: 4511768D6622BEE5 CRC64:

Query Match 89.5%; Score 263; DB 1; Length 110;
Best Local Similarity 59.3%; Pred. No. 1.9e-25;
Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNHLCGSHVEALYLVGGERGFYTPK----- 30
Db 25 FVNHLCGSHVEALYLVGGERGFYTPKSRREVERELVGGAEIGGPGAGLQPSALEL 84
QY 31 ----RGIVGCGCTSGSLYOLENYCN 52
Db 85 ALOKRGIVEGCTSGSLYOLENYCN 110

RESULT 14
INS_FELCA
ID INS_FELCA STANDARD: PRT: 51 AA.
AC P06306;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
DE INS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OC NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RA MEDLINE=86214076; PubMed=3518635;
RA Hallden G., Gavelin G., Matt V., Jorcnvall H.;
RT "Characterization of cat insulin."
RL Arch. Biochem. Biophys. 247:20-27(1986).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A01588; INCT.
DR HSSP: P01317; IAPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR PRINTS: PR00277; INSULIN.

```

```

DR SMART: SM00078; ILGF: 1.
DR PROSITE: PS00262; INSULIN: 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 36 INSULIN B CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA: 5745 MW: 9007509660A7500 PROL:
Query Match 39.3% Score 262.5; DB 1; Length 51;
Best Local Similarity 90.4%; Pred. No. 1.1e-25;
Matches 47; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
QY 1 FVNHQCGSHLVKALYVCGGEGFFTPKTKGIVEQCCTSCSLYQLENYCN 52
III|IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 1 FVNHQCGSHLVKALYVCGGEGFFTPKA-GIVEQCTASVCSLYQLENYCN 51
Query Match 89.1% Score 262; DB 1; Length 110;
Best Local Similarity 59.3%; Pred. No. 2.6e-25;
Matches 51; Conservative 0; Mismatches 1; Indels 34; Gaps 1;
QY 1 FVNHQCGSHLVKALYVCGGEGFFTPKTKGIVEQCCTSCSLYQLENYCN 52
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 25 FVNHQCGSHLVKALYVCGGEGFFTPKARREVEDLVQYRDVELAGPCEGGLQPLAEG 84
QY 41 ----RGIVEQCCTSCSLYQLENYCN 52
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 85 ALQKRGIVEQCCTSCSLYQLENYCN 110
Search completed: September 15, 2003, 12:01:30
Job time : 4.84588 secs

```

RESULT 15

```

INS_CANFA
ID INS_CANFA STANDARD: PRT: 110 AA.
AC P01321;
DT 21-JUL-1985 (Rel. 61, Created)
DT 21-JUL-1986 (Rel. 61, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
OS Canis familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX MCB_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84109071; PubMed=6296342;
RA KWOK S.C.M., Chan S.S., Stojner D.F.
RT "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT Coded amino acid sequence of canine proinsulin predicts an
RT additional C-peptide fragment."
RL J. Biol. Chem. 258:2357-2353(1983).
RN [2]
RP SEQUENCE OF 25-54 AND 95-110.
RX MEDLINE=66160119; PubMed=5949523;
RA Smith L.P.;
RL Am. J. Med. 40:662-665(1966);
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is provided through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: V00179; CAA23475.1;
CC PIR: A92413; IPEG.
CC HSSP: P01317; IAPR.
CC InterPro: IPR004825; Irs/IGF/relax.
CC Pfam: PF00049; Insulin; 1.
CC SMART: SM00078; ILGF: 1.
CC PROSITE: PS00262; INSULIN: 1.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24

```

GenCore version 5.1.4
Copyright (c) 1993 2003 Computer Ltd.

OM protein - protein search, using sw Model:

Run on: September 15, 2003, 11:54:30 : Search time 27.314 seconds
(without alignments)
669,520 Million Cells updates/sec

Title: US-09-423-100-5
Perfect score: 234
Sequence: 1 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 52

Scoring table: ELOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_23:
1: sp-archaea:
2: sp-bacteria:
3: sp-fungi:
4: sp-human:
5: sp-invertebrate:
6: sp-mammal:
7: sp-misc:
8: sp-organelle:
9: sp-phage:
10: sp-p-anti:
11: sp-rodent:
12: sp-virus:
13: sp-vertebrate:
14: sp-undclassified:
15: sp-viruses:
16: sp-bacteriopl:
17: sp-archaea:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	90.8	110	Q8HXV2	Q8HXV2 Pongo pygmaeus (Orangutan).
2	265	89.5	110	Q91X13	Q91X13 Homo sapiens (Human).
3	251	85.4	110	Q8WKK6	Q8WKK6 Homo sapiens (Human).
4	219.5	74.7	106	Q918Q7	Q918Q7 Homo sapiens (Human).
5	201.5	68.5	110	Q98TA6	Q98TA6 Pteropus vampyrus (Flying fox).
6	195.5	66.5	108	Q90DE5	Q90DE5 Brachycephalus excelsus (Tree toad).
7	195.5	66.5	108	Q90ZM4	Q90ZM4 Callisaurus draconoides (Spiny-tailed lizard).
8	195	66.3	113	Q98TB0	Q98TB0 Chelonia mydas (Sea turtle).
9	193.5	65.8	110	Q98ZV1	Q98ZV1 Briconia allos (Fruit fly).
10	191.5	65.1	111	Q98TA7	Q98TA7 Acanthoscytus (Weevil).
11	187.5	63.8	87	Q98TA9	Q98TA9 Acanthoscytus (Weevil).
12	186.5	63.4	108	Q98TB1	Q98TB1 Acanthoscytus (Weevil).
13	185.5	63.1	91	Q98TB2	Q98TB2 Acanthoscytus (Weevil).
14	146	49.7	65	Q8H281	Q8H281 Acrida gor (Grasshopper).
15	146	49.7	65	Q8H280	Q8H280 Pongo pygmaeus (Orangutan).
16	144	49.0	207	Q90XD0	Q90XD0 Cypripedium (Orchid).

ALIGNMENTS

RESULT 1:

Q8HXV2 PRELIMINARY: PRT: 110 AA.
ID Q8HXV2
AC Q8HXV2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin precursor.
GN INS.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI TaxID:9606;
ER 1
RF 1
RA Stead J.D.H., Jeffreys A.J.:
RT "Haplotype diversity at the insulin region."
RL Submitted (JUL-2002) to the EMBL/GenBank/CDL databases.
PR EMBL: AV137503; AAC06537; J01777
SQ SEQUENCE 110 AA, 12018 MW: 22023209452048 CRC64;

Query Match 90.8%; Score 267; DB 6; Length 110;
Best local similarity 62.5%; Prod. No. 1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 2;

CY 1 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 52
DB 25 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 110
CY 31 ----SGIVEQCCTSCSYOLENYCN 52
DB 65 SLOKRGVEQCCTSCSYOLENYCN 110

RESULT 2:

Q91X13 PRELIMINARY: PRT: 110 AA.
AC Q91X13
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

073721 tilapia sp.
Q8AV14 petromyzon
Q93607 paralichthy
Q73720 oreochromis
Q42289 oreochromis
P75824 oreochromis
Q57436 paralichthy
Q93527 paralichthy
Q8V157 acanthopagr
Q91161 oncorhynch
Q91476 salmo salar
Q91475 salmo salar
Q91231 oncorhynch
Q91162 oncorhynch
Q91230 oncorhynch
Q42336 myoxocephal
P81268 oncorhynch
Q91965 oncorhynch
Q91443 squalus aca
Q91164 brachydanic
Q42429 lates calca
Q8MJ15 sus scrofa
Q93380 melegritis g
Q91440 carassius a
Q91914 ctenopharyn
Q9MYX4 bos indicus
Q90VV9 brachydanic
Q9PWK2 carassius a
Q95SR6 megalobrama

```

DT 01-DEC-2001 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scurinae;
OC Spermophilus.
CX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
KA Tredrea M.M., Buck M.J., Gubanyioui J., Squitieri L.L., Andrews M.L.;
RT "Regulation of PDx4 expression in a bipednating mamma...";
RL Submitted (JUN-2001) to the EMBL/GenBank/CCDB databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC EMBL: AY038504; AAK72558.1; -.
DR HSSP: P01308; 11NP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12604 MW: 451176665228885 CRC64:
Query Match 89.5%; Score 263; DB 11; Length 110;
Best Local Similarity 59.3%; Pred. No. 3.6e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 4; Gaps 1.
QY 1 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
DB 25 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
QY 31 ----RGIVECCCTSGSLYLENYCN 52
DB 85 ALKRGIVGCCISGLYLENYCN 110
Query Match 74.7%; Score 219.5; DB 13; Length 106;
Best Local Similarity 49.4%; Pred. No. 3e-22;
Matches 41; Conservative 7; Mismatches 4; Indels 31; Gaps 1.
RESULT 3
ID Q8WNK6 PRELIMINARY; PRT; 110 AA.
AC Q8WNK6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
CX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
KA Okamoto S., Morimatsu M.;
RT "Cat insulin.";
RL Submitted (MAY-2000) to the EMBL/GenBank/CCDB databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AB043535; BAB84110.1; -.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12069 MW: 95F86E170C75E2A4 CRC64:
Query Match 85.4%; Score 251; DB 6; Length 110;
Best Local Similarity 55.8%; Pred. No. 1.6e-26;
Matches 48; Conservative 2; Mismatches 2; Indels 4; Gaps 1.
QY 1 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
DB 25 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.

```

```

QY 31 ----RGIVECCCTSGSLYLENYCN 52
DB 85 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
Query Match 74.7%; Score 219.5; DB 13; Length 106;
Best Local Similarity 49.4%; Pred. No. 3e-22;
Matches 41; Conservative 7; Mismatches 4; Indels 31; Gaps 1.
RESULT 4
ID Q91507 PRELIMINARY; PRT; 106 AA.
AC Q91507;
DT 01-DEC-2001 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana;
CX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20362507; PubMed=10818274;
KA Irwin D.M., Sivarajah P.;
RT "Proinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
proinsulin processing.";
ML Comp. Biochem. Physiol. 125B:405-410(2000).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC EMBL: AF227167; AAF87285.1; -.
CC HSSP: P01315; 1SDR.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00049; Insulin; 1.
CC PRINTS: PR00277; INSULIN.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 106 AA: 12183 MW: 3AB7CEC70217F92 CRC64:
Query Match 74.7%; Score 219.5; DB 13; Length 106;
Best Local Similarity 49.4%; Pred. No. 3e-22;
Matches 41; Conservative 7; Mismatches 4; Indels 31; Gaps 1.
QY 1 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
DB 24 FVNHLCGSHLVEALYVCGRGGFFFTYKTR----- 110 AA.
QY 32 ----GIVECCCTSGSLYLENYCN 52
DB 84 KPGIVECCCTSGSLYLENYCN 106
Query Match 74.7%; Score 219.5; DB 13; Length 106;
Best Local Similarity 49.4%; Pred. No. 3e-22;
Matches 41; Conservative 7; Mismatches 4; Indels 31; Gaps 1.
RESULT 5
ID Q98TA8 PRELIMINARY; PRT; 110 AA.
AC Q98TA8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
CX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21203577; PubMed=11306171;
KA Al-Mahroki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid.";
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC EMBL: AF195368; AAK28712.1; -.
DR HSSP: P01308; 1HIS.

```



```
QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 26 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDVPEPDQAE 85
QY 30 ---TRGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 26 VKMKRGIVEQCCHRPNCNFDLQNYCN 112

RESULT 9
Q902Y1 ID Q902Y1 PRELIMINARY: PRT: 111 AA.
AC Q902Y1:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Preproinsulin (Fragment)
OS Hiodon alosoides (holidae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Hiodontidae; Hiodon.
OX NCBI_TaxID:54934.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF282406; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PG0277; INSULIN.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 112
FT NON_TER 112
SQ SEQUENCE 110 AA: 12345 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 65.8%; Score 193.5; DB 13; Length 110;
      Best Local Similarity 42.4%; Pred. No. 1.1e-18;
      Matches 36; Conservative 7; Mismatches 7; Indels 35; Gaps 1;

QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 26 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDVPEPDQAE 85
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 86 VKRGIVEQCCHRPNCNFDLQNYCN 110

RESULT 10
Q987A7 ID Q987A7 PRELIMINARY: PRT: 111 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Preproinsulin (Fragment)
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID:109271.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 111
FT NON_TER 111
SQ SEQUENCE 87 AA: 9874 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 63.8%; Score 187.5; DB 13; Length 87;
      Best Local Similarity 42.9%; Pred. No. 5.8e-18;
      Matches 36; Conservative 5; Mismatches 8; Indels 35; Gaps 1;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 4 QHLGSHLVEALFLVCGERGFNPDTRKRDVDSLLGLFLSPKSGPENEADVRYKEQAEVK 63
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 64 VKRGIVEQCCHRPNCNFDLQNYCN 87

RESULT 12
Q987A7 ID Q987A7 PRELIMINARY: PRT: 111 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Preproinsulin (Fragment)
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID:109271.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 111
FT NON_TER 111
SQ SEQUENCE 87 AA: 9874 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 63.8%; Score 187.5; DB 13; Length 87;
      Best Local Similarity 42.9%; Pred. No. 5.8e-18;
      Matches 36; Conservative 5; Mismatches 8; Indels 35; Gaps 1;
```

```
QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 26 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDVPEPDQAE 85
QY 30 ---TRGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 26 VKMKRGIVEQCCHRPNCNFDLQNYCN 112

RESULT 9
Q902Y1 ID Q902Y1 PRELIMINARY: PRT: 111 AA.
AC Q902Y1:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Preproinsulin (Fragment)
OS Hiodon alosoides (holidae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Hiodontidae; Hiodon.
OX NCBI_TaxID:54934.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF282406; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PG0277; INSULIN.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 112
FT NON_TER 112
SQ SEQUENCE 110 AA: 12345 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 65.1%; Score 191.5; DB 13; Length 111;
      Best Local Similarity 41.2%; Pred. No. 2.1e-18;
      Matches 35; Conservative 10; Mismatches 5; Indels 35; Gaps 1;

QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 27 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGPENEADVRYKEQAEVK 86
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 37 VKRGIVEQCCHRPNCNFDLQNYCN 111

RESULT 11
Q987A7 ID Q987A7 PRELIMINARY: PRT: 87 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Preproinsulin (Fragment)
OS Gnathonemus petersi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathonemus.
OX NCBI_TaxID:42645.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 87
FT NON_TER 87
SQ SEQUENCE 87 AA: 9874 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 63.8%; Score 187.5; DB 13; Length 87;
      Best Local Similarity 42.9%; Pred. No. 5.8e-18;
      Matches 36; Conservative 5; Mismatches 8; Indels 35; Gaps 1;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
Db 4 QHLGSHLVEALFLVCGERGFNPDTRKRDVDSLLGLFLSPKSGPENEADVRYKEQAEVK 63
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
Db 64 VKRGIVEQCCHRPNCNFDLQNYCN 87

RESULT 12
Q987A7 ID Q987A7 PRELIMINARY: PRT: 111 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Preproinsulin (Fragment)
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID:109271.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE:21203577; PubMed:11306171.
RA Al-Mahrouk A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; ||||| ||
DR HSSP: PG1308; IIGF; 1.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
DR NON_TER 111
FT NON_TER 111
SQ SEQUENCE 87 AA: 9874 MW: 85525256 CRC64:
      MEDLINE:21203577; PubMed:11306171.
      Query Match 63.8%; Score 187.5; DB 13; Length 87;
      Best Local Similarity 42.9%; Pred. No. 5.8e-18;
      Matches 36; Conservative 5; Mismatches 8; Indels 35; Gaps 1;
```

```

ID Q98TB2: PRELIMINARY: PRT: 128 AA.
AC Q98TB2:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=96071;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21203577; PubMed-11306171;
RA Al-Mahrouki A.A., Irwin D.M., Graham D.C., Yanson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174(S1-S9(2001)).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 108 108
FT SEQUENCE 108 AA; 11873 MW; E426310696FBA076 CRC64;

Query Match 63.4%; Score 186.5; DB 13; Length 108;
Best Local Similarity 43.9%; Pred. No. 9.6e-18;
Matches 36; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

QY 4 QHLGGSHLVFALYLVCGRGFFYTPK-----T 40
DE 27 QHLGGSHLVFALYLVCGRGFFYTPKTRRAEDLVGQVGLGGPGAGSLOPLALEGSLQKRG 60

QY 31 RGIVSQCTICSILYOLENYCN 52
DE 27 RGIVSQCTICSILYOLENYCN 52
FT NON_TER 108 108
FT SEQUENCE 91 AA; 10100 MW; E36C882560C65039 CRC64;

ID Q98TB2: PRELIMINARY: PRT: 61 AA.
AC Q98TB2:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Ambipolites rupestris (Rock bass).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neotlostei;
OC Acanthopterygii; Acanthopterygii; Perciformes; Perciformes; Perciformes;
OC Centrarchidae; Ambloplites.
OX NCBI_TaxID=103273;
RN [1]
RP SEQUENCE FROM N.A.
RX Al-Mahrouki A.A., Irwin D.M., Yanson J.H.;
RT "Molecular cloning of preproinsulin cDNA from the rock bass."
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 91 91
FT SEQUENCE 91 AA; 10100 MW; E36C882560C65039 CRC64;

```

```

Query Match 63.4%; Score 185.5; DB 13; Length 91;
Best Local Similarity 40.9%; Pred. No. 1.1e-17;
Matches 36; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

QY 4 QHLGGSHLVFALYLVCGRGFFYTPK----- 29
DE 4 QHLGGSHLVFALYLVCGRGFFYTPKTRRAEDLVGQVGLGGPGAGSLOPLALEGSLQKRG 63
FT NON_TER 1 1
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

QY 46 -----IRGIVSQCTICSILYOLENYCN 52
DE 46 -----IRGIVSQCTICSILYOLENYCN 52
FT NON_TER 1 1
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

QY 64 MEMVKKRGIVGCGCHHPNPFQIGRYCN 91
DE 64 MEMVKKRGIVGCGCHHPNPFQIGRYCN 91
FT NON_TER 1 1
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

RESULT 14
Q98TB2: PRELIMINARY: PRT: 65 AA.
AC Q98TB2:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX O'Duigain C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092023; AAM76640.1; -.
FT NON_TER 1 1
FT NON_TER 65 65
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

Query Match 49.7%; Score 146; DB 6; Length 65;
Best Local Similarity 47.7%; Pred. No. 2e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 7 QGSHLVFALYLVCGRGFFYTPK-----RG 32
DE 7 QGSHLVFALYLVCGRGFFYTPKTRRAEDLVGQVGLGGPGAGSLOPLALEGSLQKRG 60
FT NON_TER 1 1
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

QY 43 IVEQC 37
DE 43 IVEQC 65
FT NON_TER 1 1
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

RESULT 15
Q98TB2: PRELIMINARY: PRT: 65 AA.
AC Q98TB2:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX O'Duigain C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092024; AAM76641.1; -.
FT NON_TER 1 1
FT NON_TER 65 65
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

Query Match 49.7%; Score 146; DB 6; Length 65;

```

Best Local Similarity 47.78 Prod. No. 2e 12:
Matches 31: Conservative 0: Mismatches 0: Indels 34: Gaps 1:
QY 7 CGSHLVPAALVVGSGRGFFVTPKT-----EQL
Db 1 CGSHLVPAALVVGSGRGFFVTPKTREAEELVGVQVDELNGGPGAGSGLALEGSLQKAG 60
QY 33 IVROC 37
Db 61 IVROC 65

Search completed: September 15, 2003, 12:03:30
Job time : 21.5154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer 214

OK protein - protein search, using sw model

Run on: September 16, 2003, 12:31:20 : Search time 31.542 seconds
(without alignments)
536,746 Million cell updates/sec

Title: US-09-423-100-6

Perfect score: 587

Sequence: 1 MFPIILSRJEDNAPLEAHLK.....INRQVTSISLSYGLQENYCN LQ?

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 3.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1951.DAT:*
- 2: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1951.DAT:*
- 3: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1952.DAT:*
- 4: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1953.DAT:*
- 5: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1954.DAT:*
- 6: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1955.DAT:*
- 7: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1956.DAT:*
- 8: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1957.DAT:*
- 9: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1958.DAT:*
- 10: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1959.DAT:*
- 11: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1960.DAT:*
- 12: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1961.DAT:*
- 13: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1962.DAT:*
- 14: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1963.DAT:*
- 15: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1964.DAT:*
- 16: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1965.DAT:*
- 17: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1966.DAT:*
- 18: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1967.DAT:*
- 19: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1968.DAT:*
- 20: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1969.DAT:*
- 21: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1970.DAT:*
- 22: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1971.DAT:*
- 23: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1972.DAT:*
- 24: /SIDSL/qcdata/geneseq/geneseq-emb1/AA1973.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	587	100.0	107	20	AA1942860
2	555.5	94.6	150	20	AA1942861
3	315.5	51.7	116	17	AA1968897
4	304	51.8	63	15	AA1968900
5	304	51.8	117	17	AA1968896
6	302.5	51.5	137	16	AA1971692
7	299	50.9	55	15	AA1968901
8	299	50.9	56	16	AA1978665
9	299	50.9	96	15	AA1968899

10	299	50.9	96	16	AA1978662
11	299	50.9	145	16	AA1971694
12	299	50.9	146	16	AA1971695
13	294	50.1	52	20	AA1942859
14	293	49.9	57	11	AA1940562
15	289.5	49.1	360	16	AA1970556
16	287	48.9	52	12	AA1971899
17	287	48.9	65	19	AA1973663
18	287	48.9	136	17	AA1970866
19	284.5	48.5	58	17	AA1968047
20	284.5	48.5	59	17	AA1968048
21	284.5	48.5	65	17	AA1968188
22	284.5	48.5	103	10	AA1946445
23	284.5	48.5	123	18	AA19240
24	284.5	48.5	123	19	AA1969160
25	284.5	48.5	124	19	AA1978751
26	284.5	48.5	124	24	AB1982578
27	284.5	48.5	124	24	AB1985059
28	284.5	48.5	125	18	AA19242
29	284	48.4	138	10	AA1946443
30	284	48.4	138	10	AA1946444
31	284	48.4	140	16	AA1971693
32	284	48.4	140	16	AA1971690
33	283.5	48.3	53	15	AA1968883
34	283.5	48.3	53	18	AA198007
35	283.5	48.3	117	19	AA1978752
36	283.5	48.3	408	22	AA193705
37	282	48.0	94	10	AA1946444
38	282	48.0	120	18	AA19241
39	281.5	48.0	60	3	AA1920002
40	281.5	48.0	105	10	AA1946448
41	281.5	48.0	153	21	AA193589
42	281	47.9	102	20	AA1946449
43	280	47.9	104	16	AA1971684
44	280.5	47.8	55	8	AA1971019
45	280.5	47.8	89	17	AA1988179

ALIGNMENTS

RESULT 1
AA1942860
ID: AA1942860 standard; protein: 107 AA.

AC: AA1942860;

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

XX: AC

New chimeric proteins containing human growth hormone fragment, used

PT particularly for the production of human insulin -
XX
XX
XX Claim 13; Page 30; 46pp; English.
XX
CC This sequence represents a chimeric protein, hGH-mini-proinsulin.
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
CC linker (AAY42857), and a human insulin precursor comprising insulin
CC A and B chains (AAY42859). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor.
CC C-terminal Arg residue which enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphitolyis
CC and related purification steps can thus be eliminated, along with the use
CC of high concentrations of mercaptan or the use of hydrophobic absorbent
CC resins.
XX
XX Sequence 107 AA:
XX
Query Match 100.0%; Score 587; DA 20; Length 107;
Best Local Similarity 100.0%; Pred. No. 40-42;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPTIPLSRLFDNAMLRAHRLHQLAFTDYQEFEEAYIKKQKYSFLQNP----- 49
DB 1 MEPTIPLSRLFDNAMLRAHRLHQLAFTDYQEFEEAYIKKQKYSFLQNPQTSLSFSESIP 60
QY 61 LQSHLVEAHLVCGGSGFFYTKRIGVEQCTSCSLYQLENYCN 107
DB 61 LQSHLVEAHLVCGGSGFFYTKRIGVEQCTSCSLYQLENYCN 107
RESULT 2
AY42861
ID AAY42861 standard; protein: 150 AA.
XX
XX AAY42861;
XX
XX (9-JAN-2000 (first entry)
XX
XX Chimeric protein, SEQ ID 7.
XX
XX Insulin; precursor; growth hormone; chaperone; intramolecular;
XX folding; conformation; chimeric protein; cleavable; recombinant;
XX production; yield.
XX
XX Synthetic.
XX Homo sapiens.
XX WO9950302-A1.
XX
XX 07-OCT-1999.
XX
XX 31-MAR-1998; 98WO-CN00052.
XX
XX 31-MAR-1998; 98WO-CN00052.
XX
XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX
XX Gan Z;
XX
XX WPI: 1999-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used

PT particularly for the production of human insulin -
XX
XX
XX Claim 14; Page 30-31; 46pp; English.
XX
CC This sequence represents a chimeric protein, which contains an
CC N-terminal fragment of human growth hormone (hGH) of the sequence given
CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
CC precursor comprising insulin A and B chains (AAY42859). The hGH portion
CC of the chimeric protein acts as an intramolecular chaperone (IMC) for
CC the insulin precursor, enabling it to fold correctly. The cleavable
CC peptide linker has a C-terminal Arg residue which enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphitolyis
CC and related purification steps can thus be eliminated, along with the use
CC of high concentrations of mercaptan or the use of hydrophobic absorbent
CC resins.
XX
XX Sequence 150 AA:
XX
Query Match 94.6%; Score 555.5; DA 20; Length 150;
Best Local Similarity 71.3%; Pred. No. 2.4e-39;
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;
QY 1 MEPTIPLSRLFDNAMLRAHRLHQLAFTDYQEFEEAYIKKQKYSFLQNP----- 49
DB 1 MEPTIPLSRLFDNAMLRAHRLHQLAFTDYQEFEEAYIKKQKYSFLQNPQTSLSFSESIP 60
QY 50 -----LGTGPRFVNOHLOGSHLVEAHLVCGGR 77
DB 61 TFSNKEEQKSNELLRTSIILOSLEPVCLGTGPRFVNOHLOGSHLVEAHLVCGGR 120
QY 78 GFFYTPKRGIVGVEQCTSCSLYQLENYCN 107
DB 121 GFFYTPKRGIVGVEQCTSCSLYQLENYCN 150
RESULT 3
AAR96897
ID AAR96897 standard; protein: 115 AA.
XX
XX AAR96897;
XX
XX 03-FEB-1997 (first entry)
XX
XX SOD-proinsulin hybrid polypeptide.
XX
XX Insulin; proinsulin; hybrid polypeptide; protein folding;
XX enzymatic cleavage; cyanogen bromide; sulphitolyis.
XX
XX Homo sapiens.
XX
XX WO9620724-A1.
XX
XX 11-JUL-1996.
XX
XX 29-DEC-1994; 94WO-US13268.
XX
XX 29-DEC-1994; 94WO-US13268.
XX
XX 10-JAN-1995; 95ZA-0000142.
XX
XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
XX Gcreeki M, Hartman JR, Mendelovitz S;
XX

DR WPI: 1996-333766/33.
 DR N-PSDB: AAT34670.
 XX Recombinant insulin prodn. by correctly folding the insulin hybrid
 PT polypeptide - then enzymatic cleavage of folded product, does not
 PT require sulphate protection of SH nor use of cyanogen bromide.
 XX Example 1b: Figure 1: 69pp; English.
 XX A new method for the production of recombinant human insulin
 CC comprises folding a hybrid polypeptide comprising proinsulin under
 CC conditions that permit correct disulphide bond formation and
 CC subjecting that folded protein to enzymatic cleavage. The insulin
 CC produced can then be purified. This sequence is a SOD-insulin B
 CC chain-Arg-insulin A chain hybrid polypeptide and is encoded by
 CC the plasmid construct pBAST-1A7. Transformation of the proper
 CC E.coli host cells with pBAST-1A7 results in the efficient
 CC expression of the proinsulin hybrid polypeptide, useful for human
 CC insulin production. The method produces recombinant human insulin
 CC identical to the natural hormone. Hazardous and cumbersome
 CC procedures involving cyanogen bromide and sulphitolysis to protect
 CC SH groups are avoided since the entire hybrid polypeptide folds
 CC efficiently to the native structure even with the leader attached
 CC and Cys unprotected.
 XX
 SQ Sequence 116 AA:
 Query Match 53.7%; Score 315.5; DB 17; Length 116;
 Best Local Similarity 85.3%; Pred. No. 2.5e-19;
 Matches 58; Conservative 2; Mismatches 5; Indels 3; Gaps 1;
 QY 43 YSEFIONPLOT--GPRFVNOHLCGSHLVEALYVCGSGRFFYTPKTRGIVEQCSTSCS 99
 Db 49 HFGDNTAGTAGAGPRFVNOHLCGSHLVEALYVCGSGRFFYTPKTRGIVEQCSTSCS 106
 QY 100 YQLENVCN 107
 Db 109 YQLENVCN 115
 RESULT 4
 AAR68900
 ID AAR68900 standard; Peptide: 63 AA.
 AC AAR68900;
 CC 25-MAR-2003 (updated);
 DT 02-MAR-1995 (first entry)
 XX Human pro-insulin 4.
 XX Pro-insulin; A chain; B chain; C chain; disulphide;
 KW mercaptan; chaotropic agent.
 XX Homo sapiens.
 CS EP600372-A1.
 PN 08-JUN-1994.
 XX 25-NOV-1993; 93EP-0118593.
 XX 02-DEC-1992; 92SB-424042C.
 XX (FARH) HOECHST AG.
 XX Gerl M, Ludwig J, Obermeier R, Sabot W;
 XX WPI: 1994-177718/22.
 XX Prodn. of pro-insulin with correct disulphide bridges - by
 PT treating recombinant precursor protein with mercaptan in alkali
 PT and in presence of chaotropic agent, then isolation on

PI Hydrophobic resin
 XX Disclosure: Page 11-12; 15pp; German.
 XX pro-insulin is produced by treating recombinant precursor protein
 CC with a mercaptan to provide 2-10 SH residues per Cys residue, in
 CC presence of a chaotropic agent and in aq. medium of pH 10-11.
 CC treating the prod. with 3-50 g hydrophobic adsorber resin per l aq.
 CC medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
 CC desorbing the pro-insulin. This method produces pro-insulin with
 CC correctly bonded Cys bridges. Compared with known methods it
 CC involves fewer stages (esp. no sulphitolysis or cyanogen bromide
 CC cleavage) and overall losses during purification are reduced, i.e.
 CC the process is quicker and gives better yields.
 CC Sequences of insulin chain A, B and C are given in AAR68895-97.
 CC Sequences of pro-insulin 1-4 are given in AAR68898-901.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 63 AA:
 Query Match 51.8%; Score 304; DB 15; Length 63;
 Best Local Similarity 94.7%; Pred. No. 1.3e-18;
 Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 51 GTGPRFVNOHLCGSHLVEALYVCGSGRFFYTPKTRGIVEQCSTSCSLYQLENVCN 107
 Db 7 GNSARFVNOHLCGSHLVEALYVCGSGRFFYTPKTRGIVEQCSTSCSLYQLENVCN 63
 RESULT 5
 AAR68896
 ID AAR68896 standard; Protein: 117 AA.
 AC AAR68896;
 XX 03-FEB-1997 (first entry)
 DT SOD-proinsulin hybrid polypeptide.
 XX Insulin; Proinsulin; hybrid polypeptide; protein folding;
 KW enzymatic cleavage; cyanogen bromide; sulphitolysis.
 XX Homo sapiens.
 CS WO5620724-A1.
 PN 11-JUL-1996.
 XX 29-DEC-1994; 94WO-US33268.
 XX 25-DEC-1994; 94WO-US33248.
 XX 10-JAN-1995; 95ZA-0000142.
 XX (R10T-) BIO-TECHNOLOGY GENERAL CORP.
 XX Gurecki M, Hartman JP, Mendelovitz S;
 XX WPI: 1996-333766/33.
 XX N-PSDB: AAT34669.
 XX Recombinant insulin prodn. by correctly folding pro-insulin hybrid
 PT polypeptide - then enzymatic cleavage of folded product, does not
 PT require sulphate protection of SH nor use of cyanogen bromide
 XX Example 1a: Figure 6; 69pp; English.
 XX A new method for the production of recombinant human insulin
 CC comprises folding a hybrid polypeptide comprising proinsulin under
 CC conditions that permit correct disulphide bond formation and
 CC subjecting that folded protein to enzymatic cleavage. The insulin
 CC produced can then be purified. This sequence is a SOD-insulin B
 CC chain-Lys-Arg-insulin A chain hybrid polypeptide and is encoded by
 CC the plasmid construct pBAST-R. Transformation of the proper E.coli

CC host cells with pBAST-R results in the efficient expression of the
 CC proinsulin hybrid polypeptide, useful for human insulin production.
 CC The method produces recombinant human insulin identical to the
 CC natural hormone. Hazardous and cumbersome procedures involving
 CC cyanogen bromide and sulfitolysis to protect SH groups are avoided
 CC since the entire hybrid polypeptide folds efficiently to the native
 CC structure even with the leader attached and Cys unprotected.

XX Sequence 117 AA;
 SQ

Query Match 51.8%; Score 304; DB 17; Length 117;
 Best Local Similarity 82.6%; Pred. No. 2,3e-18;
 Matches 57; Conservative 3; Mismatches 5; Indels 4; Gaps 2;

QY 43 YSFIONPIGT--GPRFVNOHLCGSHLVEALYVCGRGFFYTPKTRGIVLEOCCTSIOS 56
 DB 45 HEFGNTEAGSTAGPRFVNOHLCGSHLVEALYVCGRGFFYTPKTRGIVLEOCCTSIOS 108

QY 99 LYOLENYCN 127
 DB 109 LYOLENYCN 117

RESULT 6
 AAR71692
 ID AAR71692 standard; Protein: 117 AA.
 XX
 AC AAR71692;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-NOV-1995 (first entry)
 XX
 DE Mating factor alpha 1-Insulin precursor Arg831.
 XX
 KW Human insulin precursor Arg831; diabetes; zinc ion complex.
 KW Mating factor alpha 1.
 XX
 OS Homo sapiens.
 XX
 XX Key location/Qualifiers
 FT Protein 1..85
 FT /Label: mating factor alpha-1
 FT Peptide 86..116
 FT /Label: B-chain
 FT Peptide 117..137
 FT /Label: A-chain
 XX
 PN W09507931-A1.
 XX
 PD 23-MAR-1995.
 XX
 PF 16-SEP-1994; 94MO-DK00347.
 XX
 PR 17-SEP-1993; 93DK-0001047.
 PR 02-FEB-1994; 94US-0190823.
 XX
 PA (NOVC) NOVO-NORDISK AS.
 XX
 XX Andersen AS, Halstrom JH, Havelund S, Jonassen J;
 PI Markussen J;
 XX
 XX WPI: 1995-131314/17.
 DR N-PSDB; AAQ86425.
 XX
 PT Acylated insulin deriv. which may be present as a zinc ion
 PT complex - is used to treat diabetes and is rapid acting
 XX
 PS Example 5; Page 78; 100pp; English.
 XX
 CC AAQ86425 encodes AAR71692 mating factor alpha 1-insulin precursor
 CC Arg831. Arg831 comprises the B and A chains of a claimed human
 CC insulin derivative. In the final claimed compsr. they are covalently
 CC connected via disulphide bonds between Cys residues A7/A7 and

CC A20/B19. The derivative, which may be present as a zinc ion
 CC complex, can be used as a fast action treatment for diabetes.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 137 AA;
 Query Match 51.5%; Score 302.5; DB 16; Length 137;
 Best Local Similarity 50.0%; Pred. No. 3.6e-18;
 Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4;

QY 2 FPIPLSKLFUNALRAHRLHQLAFDTYDEFFEAYIPKEU--KYSFIO-----N 48
 DB 3 FPSI-----FIAVLFAASSALAAPVNTTTEDETAQIPAFAVIGYSLGDFDVAVLPSN 57

QY 49 PIGTG-----PRFVNOHLCGSHLVEALYVCGRGFFYTPKTRG 87
 DB 58 STNNGELFINTFIASIAAKEEGVSMKRFVNOHLCGSHLVEALYVCGRGFFYTPKTRG 117

QY 88 IVEOCCTSIOSLYOLENYCN 107
 DB 118 IVEOCCTSIOSLYOLENYCN 137

RESULT 7
 AAR68901
 ID AAR68901 standard; peptide: 56 AA.
 XX
 AC AAR68901;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-MAR-1995 (first entry)
 XX
 DE Human pro-insulin 3.
 XX
 KW Pro-insulin; A-chain; B-chain; C-chain; disulphide;
 KW mercaptan; chaotropic agent.
 XX
 OS Homo sapiens.
 XX
 PN EP606372-A1.
 XX
 DE 08-JUN-1994.
 XX
 PF 25-NOV-1993; 93EP-0118993.
 XX
 PR 02-DEC-1992; 92DE-4240420.
 XX
 PA (FARR) HCECHST AG.
 XX
 PI Gerl M, Ludwig J, Obermeier R, Sabel W;
 XX
 DR WPI: 1994-177719/22.
 XX
 XX prodn. of pro-insulin with correct disulphide bridges - by
 XX treating recombinant precursor protein with mercaptan in alkali
 XX and in presence of chaotropic agent, then isolation on
 XX hydrophobic resin
 PS Disclosure; Page 12; 15pp; German.
 CC
 CC Pro-insulin is produced by treating recombinant precursor protein
 CC with a mercaptan to provide 2-10 SH residues per Cys residue, in
 CC presence of a chaotropic agent and in aq. medium of pH 10-11,
 CC treating the prod. with 3-50 g hydrophobic adsorber resin per 1 aq.
 CC medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
 CC desorbing the pro-insulin. This method produces pro-insulin with
 CC correctly bonded Cys bridges. Compared with known methods it
 CC involves fewer stages (esp. no sulfitolysis or cyanogen bromide
 CC cleavage) and overall losses during purification are reduced, i.e.
 CC the process is quicker and gives better yields.
 CC Sequences of insulin chain A, B and C are given in AAR68895-97.
 CC Sequences of pro-insulin 1-4 are given in AAR68898-901.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 56 AA:
 Query Match 50.9%; Score 299; DB 15; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGAGSHLVEALYVGCGERGFYTPKTRGIVEGCTISICSLYLENYCN 107
 DB 4 REVNHGAGSHLVEALYVGCGERGFYTPKTRGIVEGCTISICSLYLENYCN 56

RESULT 8
 AAR78665 standard: peptide: 56 AA.
 XX AAR78665:
 XX 03-APR-1996 (first entry)
 XX Proinsulin sequence 3.
 XX Proinsulin: post-translational modification: recombinant production;
 XX protein folding: conformation.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX Region 1..4
 FT /label= R2
 FT /note= "a peptide of 4 amino acids"
 FT Peptide 5..34
 FT /label= R1-(R2-R25) Y
 FT /note= "human insulin B-chain"
 FT Region 35
 FT /label= X
 FT Peptide 36..56
 FT /label= Gly-(A2-A20)-R3
 FT /note= "human insulin A-chain"
 XX EP668292-A2.
 XX 23-AUG-1995.
 XX 09-FEB-1995: 95EP-Q101748.
 XX 18-FEB-1994: 94DE-4405179.
 XX (FARH) HOFCHST AG.
 XX Gerl M, Ludwig J, Obermeier E, Sabel W;
 XX WPI: 1995-284754/38.
 XX Isolation of insulin that is correctly post translationally
 XX processed - by reacting proinsulin with a mercaptan in the presence
 XX of a chaotropic agent and purification after absorption to hydrophobic
 XX resin
 XX Example 2; Page 13; 16pp; German.
 XX The present sequence is an example of a proinsulin molecule prepared
 XX to the general formula R2-R1-(R2-R25) Y-X-Gly-(A2-A20)-R3 (1). In
 XX formula (1), X = Lys, Arg or a peptide of 2-15 amino acids, proline,
 XX Lys or Arg at the N- and C-termini; Y = a natural amino acid; R1
 XX enc or a bond; R2 = H, Arg, Lys, a peptide of 2-45 amino acids
 XX containing Arg or Lys at the N- and C-termini; R3 = a natural amino
 XX acid; (A2-A20) and (R2-R25) are the insulin A- and B-chain sequences
 XX from human or other insulin. The proinsulin molecule (produced in
 XX recombinant E. coli) is reacted with mercaptan at a ratio of 2:10 SH
 XX residues of mercaptan per Cys residue of proinsulin. The reaction
 XX takes place in the presence of a chaotropic auxiliary agent at
 XX pH 10-11 and results in proinsulin with correctly linked cystine

XX Bridges. Reaction with trypsin and opt., carboxypeptidase B yields
 XX correctly folded insulin. The insulin is isolated by absorption on
 XX a hydrophobic resin.
 XX Sequence 56 AA:
 Query Match 50.9%; Score 299; DB 16; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGAGSHLVEALYVGCGERGFYTPKTRGIVEGCTISICSLYLENYCN 107
 DB 4 REVNHGAGSHLVEALYVGCGERGFYTPKTRGIVEGCTISICSLYLENYCN 56

RESULT 9
 AAR6899 standard: peptide: 96 AA.
 XX AAR6899:
 XX AAR66699;
 XX 25-MAR-2003 (updated)
 XX 02-MAR-1995 (first entry)
 XX Human pro-insulin 2.
 XX Pro-insulin: A-chain; B-chain; C-chain; disulphide;
 XX mercaptan: chaotropic agent.
 XX Homo sapiens.
 XX EP600372-A1.
 XX 08-JUN-1994.
 XX 25-MAY-1993: 93EP-Q1-8993.
 XX 02-DEC-1992: 92DE-424C420.
 XX (FARH) HOFCHST AG.
 XX Gerl M, Ludwig J, Obermeier R, Sabel W;
 XX WPI: 1994-177718/22.
 XX Prodn. of pro-insulin with correct disulphide bridges - by
 XX treating recombinant precursor protein with mercaptan in alkali
 XX and in presence of chaotropic agent, then isolation on
 XX hydrophobic resin
 XX Disclosure: Page 11; 15pp; German.
 XX Pro-insulin is produced by treating recombinant precursor protein
 XX with a mercaptan to provide 2:10 SH residues per Cys residue. In
 XX presence of a chaotropic agent and in aq. medium of pH 10-11,
 XX treating the prod. with 3-50 g hydrophobic adsorber resin per 1 aq.
 XX medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
 XX desorbing the pro-insulin. This method produces pro-insulin with
 XX correctly bonded Cys bridges. Compared with known methods it
 XX involves fewer stages (esp. no sulphitolysis or cyanogen bromide
 XX cleavage) and overall losses during purification are reduced, i.e.
 XX the process is quicker and gives better yields.
 XX Sequences of insulin chain A, B and C are given in AAR6895-97.
 XX Sequences of pro-insulin 1-4 are given in AAR6898-901.
 XX Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 96 AA:
 Query Match 50.9%; Score 299; DB 15; Length 96;
 Best Local Similarity 100.0%; Pred. No. 5.1e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGAGSHLVEALYVGCGERGFYTPKTRGIVEGCTISICSLYLENYCN 107

Dd
95 RVNVQHGAGSELVALYVGGHPGFYEKTRSLVFQNTSTLSLGLSEDE IAEI HLL LHL LLLLH LHH LHH LHH LHH LHH
RESULT 12
AAR71695
ID AAR71695 standard; Protein; 146 AA.
XX XX AAK71685;
DT 25-MAR-2003 (updated)
XT 20-NOV-1995 (first entry)
XX XX Mating factor alpha 1-insulin precursor ArgB1, ARG831 N-terminal; Human insulin precursor ArgA2, ARG831; diabetes; Zinc ion complexed mating factor alpha 1; N-terminal EEAAEAER. KW KW KK KK XX XX Homo sapiens.
XX Key Location/Occalliers
FH Protein: 1..85
FT /label mating factor alpha-1
FT Peptide 86..94
FT /label- N-terminal peptide
FT Peptide 95..125
FT FT /label= B-chain
FT Peptide 126..146
FT FT /label= A-chain
XX WO9507931-A1.
PN XX
PD 23-MAR-1995.
XX XX
PP 16-SEP-1994; 94WO-DK00347.
PR 17-SEP-1993; 93DK-O001044.
XX 02-FEB-1994; 94CS-0190829.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen AS, Halstrom JB, Havelund S, Joransen F;
PT Markussen J;
DR WFI: 1995-131314/17.
PS N-PSDB: AAA086432.
XX Acylated insulin deriv. which may be present as a zinc ion complex - is used to treat diabetes and is rapidly acting.
PS Example 6; Page 85; 160pp; English.
XX AAC066432 encodes AAR71695 mating factor alpha 1-insulin precursor ArgB1, ARG831 N-terminal EEAAEAER. The insulin precursor comprised the B and A chains of a claimed Human Insulin derivative prepared by the N-terminal amino acids EEAAEAER. In the final claimed compo they are covalently connected via disulfide bonds between Cys residues A7/A7 and A20/H19. The derivative, which may be present, as a zinc ion complex, can be used as a fast action treatment for diabetes.
CC updated on 25-MAR-2003 to correct PN field.)
XX XX Sequence 146 AA:
SQ

```

RESULT 13
AAV4285
ID  AAV42859 standard; protein; 52 AA.
XX
XX  AAV42859;
XX
XX  19-JAN-2000 (first entry)
XX
XX  Human insulin precursor, SEQ ID 5.
XX
XX  Insulin; precursor; growth hormone; chaperone; intramolecular;
XX  folding; conformation; chimeric protein; cleavable; recombinant;
XX  production; yield.
XX
XX  Homo sapiens.
XX
XX  K05950302-A1.
XX
XX  07-OCT-1999.
XX
XX  31-MAR-1998; 98WO-CN00652.
XX
XX  31-MAR-1998; 98WO-CN00652.
XX
XX  (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX
XX  Gen 2;
XX
XX  WP1; 1999-610839/52.
XX
XX  New chimeric proteins containing human growth hormone fragment, used
XX  particularly for the production of human insulin
XX
XX  Claim 12; Page 29-30; 46pp; English.
XX
XX  This sequence represents a human insulin precursor comprising insulin
XX  A and B chains. This insulin precursor is a component of the chimeric
XX  proteins hGH-mim1-proinsulin (AAV42860) and the chimeric protein
XX  given in AAV42861. These chimeric proteins additionally contain an
XX  N-terminal fragment of human growth hormone (hGH) and a cleavable
XX  peptide linker (AAV42857). The hGH portion of the chimeric protein acts
XX  as an intramolecular chaperone (IMC) for the insulin precursor,
XX  enabling it to fold correctly. The cleavable peptide linker has a
XX  C-terminal Arg residue which enables the hGH portion of the
XX  chimeric protein to be removed after folding has taken place. Production
XX  of recombinant human insulin via an hGH-proinsulin chimeric protein can
XX  provide human insulin with correctly linked cysteine bridges with
XX  fewer necessary procedural steps, and hence resulting in a higher yield
XX  of human insulin. The IMC sequences not only protect insulin sequences
XX  from intracellular degradation by a microorganism host, but also promote
XX  the folding of the fused insulin precursor, facilitate the solubility of
XX  the fusion protein and decrease the intermolecular interactions among
XX  the fusion proteins, thus allowing folding of the fused insulin precursor
XX  at commercially useful high concentrations. The procedural steps of
XX  cyanogen bromide cleavage, oxidative sulfitolysis and related
XX  purification steps can thus be eliminated, along with the use of high
XX  concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
XX  Sequence 52 AA:
XX
XX  Query Match 50.1%; Score 294; DB 20; Length 52;
XX  Best Local Similarity 100.0%; Prod. No. 7.8e-18;
XX  Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps
XX
XX  56 FVNHQLCGSHLVEALYLVGGERGFYTPKINGIVGEQCTCTICS:VQLENYC 107
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  1 FVNHQLCGSHLVEALYLVGGERGFYTPKTRGIVGEQCTCS:VQLENYC 52
XX
RESULT 14
AAV4582
ID  AARG4582 standard; protein; 57 AA.
XX

```


GenData version 5.1.6
Copyright (c) 1993 - 2003 Genapan Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:37:10 Search time 13.7 seconds
(without alignments)
265,217 Million cell updates/sec

Title: US-09-423-100-6
Perfect score: 587
Sequence: 1 MPP:PISSK.FDNAM:BAKK.....IVGQNTSISQVLDNYVH IOT

Scoring table: BLSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310856 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.ppt:*
2: /cgn2_6/ptodata/1/iaa/5A_COMB.ppt:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.ppt:*
4: /cgn2_6/ptodata/1/iaa/5A_COMB.ppt:*
5: /cgn2_6/ptodata/1/iaa/PCIOUS_COMB.ppt:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.ppt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	304	51.8	63	1	US-08-160-376A-6
2	302.5	51.5	137	3	US-08-400-256-39
3	302.5	51.5	137	3	US-08-975-365-29
4	299.5	51.0	66	1	US-08-293-060P-6
5	299	50.9	56	1	US-08-160-376A-7
6	299	50.9	56	1	US-08-389-487-11
7	299	50.9	96	1	US-08-160-376A-5
8	299	50.9	96	1	US-08-389-487-8
9	299	50.9	145	3	US-08-400-256-45
10	299	50.9	145	3	US-08-975-365-45
11	299	50.9	146	1	US-08-400-256-48
12	299	50.9	146	3	US-08-975-365-48
13	293	49.3	57	1	US-08-330-733A-44
14	287	48.9	65	3	US-08-900-574-7
15	286.5	48.6	66	3	US-08-900-574-6
16	286	48.7	67	3	US-08-900-574-7
17	284.5	48.5	65	1	US-08-468-674P-71
18	284.5	48.5	65	1	US-08-780-571-71
19	284.5	48.5	124	3	US-08-012-669F-18
20	284	48.4	138	3	US-08-932-582-19
21	284	48.4	140	1	US-08-400-256-33
22	284	48.4	140	1	US-08-400-256-42
23	284	48.4	140	3	US-08-975-365-33
24	284	48.4	140	3	US-08-975-365-42
25	283.5	48.3	53	1	US-08-233-617-4
26	283.5	48.3	53	4	US-08-981-988A-42
27	283.5	48.3	117	3	US-09-012-669F-37

28	281	47.9	104	1	US-08-400-256-15	Sequence 15, Appl
29	281	47.9	104	3	US-08-975-365-15	Sequence 15, Appl
30	280.5	47.8	89	1	US-08-468-674B-41	Sequence 41, Appl
31	280.5	47.8	89	1	US-08-780-571-41	Sequence 41, Appl
32	280.5	47.8	91	1	US-08-468-674B-45	Sequence 45, Appl
33	280.5	47.8	91	1	US-08-780-571-45	Sequence 45, Appl
34	280.5	47.8	124	1	US-08-446-646-3	Sequence 3, Appl
35	280.5	47.8	167	1	US-07-918-953-8	Sequence 8, Appl
36	280.5	47.6	167	1	US-08-081-661-8	Sequence 8, Appl
37	278.5	47.4	51	4	US-09-477-924-3	Sequence 3, Appl
38	278.5	47.4	51	4	US-09-723-881-3	Sequence 3, Appl
39	278.5	47.4	51	4	US-09-723-896-3	Sequence 63, Appl
40	278	47.4	117	4	US-09-280-630-63	Sequence 3, Appl
41	277.5	47.3	53	1	US-08-233-617-3	Sequence 3, Appl
42	277	47.2	96	2	US-09-134-836-4	Sequence 4, Appl
43	277	47.2	96	4	US-09-386-303A-4	Sequence 4, Appl
44	277	47.2	97	1	US-08-160-376A-4	Sequence 4, Appl
45	277	47.2	102	1	US-08-400-256-36	Sequence 36, Appl

ALIGNMENTS

RESULT: 1
US-08-160-376A-6
Sequence 6, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Protein
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cystine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoui, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160.376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
EXOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: ROE 92/P 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant

Query Match 53.8%; Score 304; DB 1; Length 63;
Best Local Similarity 94.7%; Pred. No. 7.3e-29;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```

TELEPHONE: (908) 231-4079
TELEX: (908) 231-2256
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
CS-08-160-478A-7

Query Match      50.9%   Score 299; DB 1; Length 56;
Best Local Similarity    100.0%; Pred.No. 2,4e-28;
Matches     53; Conservative       0; Mismatches     0; Indels        0; Gaps         0;

QY          55 RFVNOHLCGSHLVALYLVCGERGFYPYTKTRGVIVEQCCTSCSYQLENYCYN 107
              ||||| |||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB           4 RFVNOHLCGSHLVALYLVCGERGFYPYTKTRGVIVEQCCTSCSYQLENYCYN 56

RESULT 5
CS-08-349-447-11
Sequence 11, Application US/28389487
Patent No. 5663291
GENERAL INFORMATION:
APPLICANT: Oberreiter, Rainer
APPICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having
TITLE OF INVENTION: Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington.
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/68/389,487
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Elcaudi, Caloi P.
REGISTRATION NUMBER: 32,220
REFERENCE/SOCKET NUMBER: 02481.1424-CO600
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amiro acid
STEREONESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-369-487-11

Query Match      50.9%   Score 299; DB 1; Length 56;
Best Local Similarity    100.0%; Pred.No. 2,4e-28;
Matches     53; Conservative       0; Mismatches     0; Indels        0; Gaps         0;

QY          55 RFVNOHLCGSHLVALYLVCGERGFYPYTKTRGVIVEQCCTSCSYQLENYCYN 107
              ||||| |||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB           4 RFVNOHLCGSHLVALYLVCGERGFYPYTKTRGVIVEQCCTSCSYQLENYCYN 56

RESULT 7
```

US-08-160-376A-5
Sequence 5, Application US/08161376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Proteinolins Possessing Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049E/P.O. Box 2560
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160.376A
FILING DATE: December 1, 1993
CLASSIFICATION: 533
PRIORITY APPLICATION DATA:
APPLYING APPLICATION NUMBER: GE P 4249420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurel, Esq.
REGISTRATION NUMBER: 31287
REFERENCE/DOCKET NUMBER: HCE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4575
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: Not relevant.

Query Match 50.9%, Score 299; DR 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.8e+28;
Matches 53; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNOHLCGSHLVALVCGERGFFTPKTRGVGGCTSSLSGLYLENYCN 107
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44 RFVNOHLCGSHLVALVCGERGFFTPKTRGVGGCTSSLSGLYLENYCN 96

RESULT 9
US-08-400-256-45
Sequence 45, Application: US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Conassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497 No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400.256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids

US-08-160-376A-5
Sequence 8, Application US/08161376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America

Query Match 50.9%, Score 299; DR 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.8e+28;
Matches 53; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNOHLCGSHLVALVCGERGFFTPKTRGVGGCTSSLSGLYLENYCN 107
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44 RFVNOHLCGSHLVALVCGERGFFTPKTRGVGGCTSSLSGLYLENYCN 96

RESULT 9
US-08-389-487-8
Sequence 8, Application: US/08161376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America


```

: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
US-08-400-256-45

Query Match      50.9%  Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 80-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 REVNHGCGSHVLEALVLCGERGFYTPKTRGIVEQCCTSCSLYQLENYCN 107
DB 93 REVNHGCGSHVLEALVLCGERGFYTPKTRGIVEQCCTSCSLYQLENYCN 145

RESULT 10
US-08-975-365-45
: Sequence 45, Application US/08975365
: Patent No. 6011007
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: APPLICANT: Markussen, Jan
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 57504970 No. 57504970disk of No. 57504970th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/400,256
: FILING DATE: 03-MAR-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985,220-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 48:
: LENGTH: 146 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-400-256-48

Query Match      50.9%  Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 80-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 REVNHGCGSHVLEALVLCGERGFYTPKTRGIVEQCCTSCSLYQLENYCN 107
DB 94 REVNHGCGSHVLEALVLCGERGFYTPKTRGIVEQCCTSCSLYQLENYCN 145

RESULT 12
US-08-975-365-45
: Sequence 48, Application US/08975365
: Patent No. 6011007
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: APPLICANT: Markussen, Jan
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 60110070 No. 60110070disk of No. 60110070th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:

```

FILING DATE: 05-MAY-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P 38 37 273.8
 FILING DATE: 03-NOV-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P 39 27 443.7
 FILING DATE: 19-AUG-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P 40 12 816.0
 FILING DATE: 21-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Kirschner Michael K.
 REGISTRATION NUMBER: 34,852
 REFERENCE/DOCKET INFORMATION: 02481-0593-02000
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 44:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 57 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 OS: 03-030-73-A-44

```

; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-28-030-73A-44

Query Match      49.9%   Score 293:   DB 1:   Length 57:
Post Local Similarity 92.2%
Matches 51: Conservative 27; Pred. No. 1.3e-27;
Mismatch 0: Indels 0; Gaps 0;

```

```

Query Match      45.9%;   Score 293;   DB 1;   Length 57;
Best Local Similarity 92.2%;   Pred. No. 1.3e-27;
Matches 51;   Conservative 2;   Mismatches 0;   Indels 0;   Gaps 0;

55 RFVNHLCGSHLVEA-YLCGSGRGFFYPKTRGIVEQCCTCSLYQLN-YNCN 107
      |||
56 RFVNHLCGSHLVEA-YLCGSGRGFFYPKTRGIVEQCCTCSLYQLN-YNCN 107
      |||
57 RFVNHLCGSHLVEA-YLCGSGRGFFYPKTRGIVEQCCTCSLYQLN-YNCN 57
      |||

```

1 5 KVN0HLCGSHVFEALYLVCGERGEFFYPKTKGIVGCCISCSLYOLN5YCN 7
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521

CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,574
FILING DATE: July 24, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: German Application No. 6221837 19630242.0
FILING DATE: July 26, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Carol P. Einaudi
REGISTRATION NUMBER: 52,220
REFERENCE/DOCKET NUMBER: 02481.1499-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400

1 INFORMATION FOR SEQ ID NO: 1:

2 SEQUENCE CHARACTERISTICS:
3 LENGTH: 65 amino acids
4 TYPE: Amino acid
5 STRANDEDNESS: Single
6 TOPOLOGY: linear
7 MOLECULE TYPE: Protein
8 ORIGINAL SOURCE:
9 ORGANISM: Escherichia coli
10 FEATURE:
11 NAME/KEY: Protein
12 LOCATION: 1..65
13 US-08-900-574-3

Query Match 48.8% Score 286.5; DB 3; Length 65;

Best Local Similarity 89.8% Pred. No. 9e-27;

Matches 53; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 51 GTGPRFVNQHLGGSHLVEALYLVCGERGFFYTPKT---RSIVQQCTSISSVQLFNYC 106

Db 7 GNSARFVNQHLGGSHLVEALYLVCGERGFFYTPKTHRRGIVEQCCTSISSVQLFNYC 65

RESULT 15

US-08-900-574-5
1 Sequence 5, Application US/08060574
2 Patent No. 6221837

3 GENERAL INFORMATION:

4 APPLICANT: Etil, Johann
5 APPLICANT: Habermann, Paul
6 APPLICANT: Geisen, Karl
7 APPLICANT: Seipke, Gerhard
8 TITLE OF INVENTION: Insulin derivatives with increased zinc
9 TITLE OF INVENTION: Binding
10 NUMBER OF SEQUENCES: 18

11 CORRESPONDENCE ADDRESS:

12 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett,
13 ADDRESSEE: & Dunner, L.L.P.
14 STREET: 1300 I Street, N.W.
15 CITY: Washington
16 STATE: District of Columbia
17 COUNTRY: U.S.A.
18 ZIP: 20005-3315

19 COMPUTER READABLE FORM:

20 MEDIUM TYPE: Floppy disk
21 COMPUTER: IBM PC compatible
22 OPERATING SYSTEM: PC-DOS/MS-DOS
23 SOFTWARE: PatentIn Release #1.0, Version #1.25 (RPO)

24 CURRENT APPLICATION DATA:

25 APPLICATION NUMBER: US/08/900,574
26 FILING DATE: July 24, 1997

27 CLASSIFICATION: 514

28 PRIOR APPLICATION DATA:

29 APPLICATION NUMBER: German Application No. 4221837 19613242.0

30 FILING DATE: July 26, 1995

31 ATTORNEY/AGENT INFORMATION:

32 NAME: Carol P. Eicard

33 REGISTRATION NUMBER: 52,320

34 REFERENCE/DOCKET NUMBER: 02481.14-99-00020

35 TELECOMMUNICATION INFORMATION:

36 TELEPHONE: (202) 408-4006

37 TELEFAX: (202) 408-4400

38 INFORMATION FOR SEQ ID NO: 5:

39 SEQUENCE CHARACTERISTICS:

40 LENGTH: 66 amino acids

41 TYPE: Amino acid

42 STRANDEDNESS: Single

43 TOPOLOGY: linear

44 MOLECULE TYPE: Protein

45 ORIGINAL SOURCE:

46 ORGANISM: Escherichia coli

47 FEATURE:
48 NAME/KEY: Protein

1 LOCATION: 1..66

2 US-08-900-574-5

3 Query Match

4 Best Local Similarity 89.8% Pred. No. 9e-27;

5 Matches 53; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 51 GTGPRFVNQHLGGSHLVEALYLVCGERGFFYTPKT---RSIVQQCTSISSVQLFNYC 106

Db 7 GNSARFVNQHLGGSHLVEALYLVCGERGFFYTPKTHRRGIVEQCCTSISSVQLFNYC 65

Search completed: September 16, 2003, 12:41:24

Run time : 18.07 secs

GenCore version: 5.1.5
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:38:55 : Search time 33.7233 seconds
(without alignments)
472.415 Million cell operations/sec

Title: US-09-423-100-6

Perfect score: 587

Sequence: 1 MFPTIP:SPFLPDNA-MLRARR.....INRGDQT:DSLSLYLENNVQ:102

Scoring table: BLASTSUM62

Searched: Gapex 13.0 , Gapex1 2.5

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 6
Maximum DB seq length: 260038000

Post-processing: Minimum Match: 0%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEK_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	107	14	US-10-054-873-6 Sequence 6, Appli
2	555.5	94.6	150	14	US-10-054-873-7 Sequence 6, Appli
3	294	50.1	52	14	US-10-054-873-5 Sequence 6, Appli
4	264.5	48.5	124	10	US-09-894-711-4 Sequence 10, Appli
5	284	48.4	138	10	US-09-894-711-5 Sequence 10, Appli
6	278.5	47.4	51	13	US-09-854-935R-5 Sequence 6, Appli
7	278.5	47.4	51	14	US-10-028-410-3 Sequence 6, Appli
8	278	47.4	117	9	US-09-280-030-6 Sequence 13, Appli
9	277	47.2	96	10	US-09-947-554-4 Sequence 4, Appli
10	275.5	46.9	124	9	US-09-736-611-12 Sequence 12, Appli
11	275.5	46.9	124	9	US-09-740-359-12 Sequence 12, Appli
12	275.5	46.9	124	10	US-09-804-711-2 Sequence 12, Appli
13	275.5	46.9	125	9	US-09-736-611-10 Sequence 10, Appli
14	275.5	46.9	125	9	US-09-740-359-10 Sequence 10, Appli
15	275.5	46.9	125	10	US-09-894-711-10 Sequence 10, Appli

16	275.5	46.9	147	9	US-09-736-611-8 Sequence 8, Appli
17	275.5	46.9	147	9	US-09-740-359-7 Sequence 7, Appli
18	274	46.7	144	9	US-09-736-611-6 Sequence 6, Appli
19	274	46.7	144	9	US-09-740-359-5 Sequence 5, Appli
20	274	46.7	146	10	US-09-894-711-5 Sequence 5, Appli
21	273	46.5	50	14	US-10-056-009A-3 Sequence 3, Appli
22	271	46.2	96	10	US-09-947-563-5 Sequence 5, Appli
23	269.5	45.9	130	9	US-09-280-030-62 Sequence 62, Appli
24	267	45.5	86	10	US-09-978-380-1 Sequence 1, Appli
25	267	45.5	86	11	US-09-958-935B-4 Sequence 4, Appli
26	267	45.5	86	14	US-10-028-410-2 Sequence 2, Appli
27	267	45.5	86	14	US-10-054-873-4 Sequence 4, Appli
28	267	45.5	110	9	US-09-255-656-125 Sequence 125, App
29	267	45.5	110	9	US-09-815-229-3 Sequence 3, Appli
30	267	45.5	110	10	US-09-504-409A-9 Sequence 9, Appli
31	267	45.5	110	12	US-09-969-749C-6 Sequence 6, Appli
32	267	45.5	110	15	US-10-338-686-1 Sequence 1, Appli
33	267	45.5	110	15	US-10-328-813-2 Sequence 2, Appli
34	260	44.3	45	14	US-10-054-873-1 Sequence 1, Appli
35	260	44.3	92	14	US-10-054-873-2 Sequence 2, Appli
36	255.5	42.5	131	11	US-09-984-010-23 Sequence 23, Appli
37	255.5	42.5	131	12	US-10-153-207-1 Sequence 1, Appli
38	255.5	42.5	191	12	US-10-400-377-1 Sequence 1, Appli
39	255.5	42.5	191	12	US-10-400-708-1 Sequence 1, Appli
40	255.5	42.5	191	12	US-10-398-148-1 Sequence 6, Appli
41	255.5	42.5	214	12	US-10-153-207-6 Sequence 9, Appli
42	255.5	42.5	217	9	US-09-929-976-5 Sequence 66, Appli
43	255.5	42.5	245	9	US-09-280-030-56 Sequence 7, Appli
44	253.5	43.2	144	10	US-09-894-711-7 Sequence 2, Appli
45	249.5	42.5	217	9	US-09-853-688-2

ALIGNMENTS

RESULT :
US-10-054-873-6
Sequence 6, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Oligomeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ionsend and Ionsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WC PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-00013005
INFORMATION FOR SEQ. IT. NO. 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid

```

: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-054-873-6

Query Match
Best Local Similarity 100.0%; Score 587; DB 14; Length 157;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY 1 MFPTIPSLRFDNAMLRAIRLHOLAFDYGEFEFAYPKCKKYSFLGNPLGTGPRFVNCH 60
DB 1 MFPTIPSLRFDNAMLRAIRLHOLAFDYGEFEFAYPKCKKYSFLGNPLGTGPRFVNCH 60

OY 61 LCGSHLVEALYVGEGGFYTPKTRGIVGECCTSCISLYOLENYCN 107
DB 1 LCGSHLVEALYVGEGGFYTPKTRGIVGECCTSCISLYOLENYCN 107

RESULT 2
US-10-054-873-7
: Sequence 7, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Ru
: TITLE OF INVENTION: Chimeric Protein Containing an
: Intramolecular Chaperone-Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/00052
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 11-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 46,946
: REFERENCE/DOCKET NUMBER: 020167-00013005
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match
Best Local Similarity 94.0%; Score 555.5; DB 14; Length 150;
Matches 107; Conservative 0; Mismatches 5; Indels 43; Gaps 1:

OY 1 MFPTIPSLRFDNAMLRAIRLHOLAFDYGEFEFAYPKCKKYSFLGNP----- 49
DB 1 MFPTIPSLRFDNAMLRAIRLHOLAFDYGEFEFAYPKCKKYSFLGNPQTSLSFSSSL 60

OY 50 -----LGIQSRFVNCHLGGSHLVEALYVCGER 77
DB 1 -----LGIQSRFVNCHLGGSHLVEALYVCGER 77

```

```

DB 61 TFSNREETQCKSNLELRISLLIQSNLEFVGLGTGRFVNOHLCGSHLVEALYVCGER 120
OY 78 GFPTPTKTRGIVGECCTSCISLYOLENYCN 107
DB 12 GFPTPTKTRGIVGECCTSCISLYOLENYCN 150

RESULT 3
US-10-054-873-5
: Sequence 5, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Ru
: TITLE OF INVENTION: Chimeric Protein Containing an
: Intramolecular Chaperone-Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/00052
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 11-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 46,946
: REFERENCE/DOCKET NUMBER: 020167-00013005
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 52 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-054-873-5

Query Match
Best Local Similarity 100.0%; Score 294; DB 14; Length 52;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY 56 FVNQHLGGSHLVEALYVCGERGFYTPKTRGIVGECCTSCISLYOLENYCN 107
DB 1 FVNQHLGGSHLVEALYVCGERGFYTPKTRGIVGECCTSCISLYOLENYCN 52

RESULT 4
US-09-894-711-18
: Sequence 18, Application US/09894711
: Patent No. US20020137144A1
: GENERAL INFORMATION:
: APPLICANT: Kjeldsen, Thomas Borglum
: APPLICANT: Ludvigsen, Svend
: TITLE OF INVENTION: Method for making insulin precursors and
: insulin precursor analogues having improved fermentation
: TITLE OF INVENTION: Insulin precursor analogues having improved fermentation
: TITLE OF INVENTION: Yield in yeast
: FILE REFERENCE: 6148, 400-US
: CURRENT APPLICATION NUMBER: US/09/894,711
: CURRENT FILING DATE: 2001-06-28

```



```

Query Match      45.9%; Score 275.5; DA 3; Length 124;
Best Local Similarity 90.9%; Pred. No. 2,1e-25;
Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 54 PFVNQHLCGSHLVEALYLVCGERGFFYTPKT RGLVGGCTISCSLYOLENYCN 107
DB 70 PFVNQHLCGSHLVEALYLVCGERGFFYTPKRWGIVFOCKTISCSLYOLENYCN 124

RESULT 11
US-09-740-359-12
: Sequence 12, Application US/09740359
: Patent No. US20050041787A1
: GENERAL INFORMATION:
: APPLICANT: Kjolseter, Thomas Borjilum
: APPLICANT: Lovdqvist, Svend
: TITLE OF INVENTION: Method for making insulin precursors and
: TITLE OF INVENTION: Insulin precursor analogues having improved formulation
: TITLE OF INVENTION: yield in yeast
: FILE REFERENCE: 6148, 200-US
: CURRENT APPLICATION NUMBER: US/09/740,359
: CURRENT FILING DATE: 2000-12-19
: PRIOR APPLICATION NUMBER: PA 2000 00443
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: PA 1999 01869
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: 60/211,081
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: 60/281,450
: PRIOR FILING DATE: 2000-02-10
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12
: LENGTH: 124
: TYPE: PRT
: ORGANISM: TA57 leader fused with N-terminally extended...
US-09-740-359-12

Query Match      46.9%; Score 275.5; DA 3; Length 124;
Best Local Similarity 90.9%; Pred. No. 2,1e-25;
Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 54 PFVNQHLCGSHLVEALYLVCGERGFFYTPKT RGLVGGCTISCSLYOLENYCN 107
DB 70 PFVNQHLCGSHLVEALYLVCGERGFFYTPKDKGCVRGVCTISCSLYOLENYCN 124

RESULT 12
US-09-694-711-12
: Sequence 12, Application US/09694711
: Patent No. US20020137144A1
: GENERAL INFORMATION:
: APPLICANT: Kjolseter, Thomas Borjilum
: APPLICANT: Lovdqvist, Svend
: TITLE OF INVENTION: Method for making insulin precursors and
: TITLE OF INVENTION: Insulin precursor analogues having improved formulation
: TITLE OF INVENTION: yield in yeast
: FILE REFERENCE: 6148, 430-US
: CURRENT APPLICATION NUMBER: US/09/694,711
: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: PA 2000 00443
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: PA 1999 01869
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: 60/211,081
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: 60/281,450
: PRIOR FILING DATE: 2000-02-10
: PRIOR APPLICATION NUMBER: 09/740,359
: PRIOR FILING DATE: 2000-12-19
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12

```

```

1 LENGTH: 124
2 TYPE: PRT
3 ORGANISM: Artificial Sequence
4 FEATURE:
5 OTHER INFORMATION: Synthetic
6
7 US-09-894-711-12
8
9 Query Match 46.9%; Score 275.5; DB 10; Length 124;
10 Best Local Similarity 90.9%; Pred. No. 2,1e-25;
11 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
12
13
14 54 PFVNOHLCGSHLVALYIVCGERGFFVTPKT-RGIVEQCCTSIICSLYOLENYCN 107
15 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 70 PFVNOHLCGSHLVALYIVCGERGFFVTDKKGIVEQCCTSIICSLYOLENYCN 124
17
18
19 RESULT 13
20 US-09-740-611-10
21 Sequence 10, Application US/09736611
22 Patient No. US2C010023069A1
23 GENERAL INFORMATION:
24 APPLICANT: Kjedtsen, Thomas
25 APPLICANT: Ludvigsen, Svend
26 APPLICANT: Kaarshoim, Niels
27 TITLE OF INVENTION: Method For Making Insulin Precursors and
28 TITLE OF INVENTION: Insulin Precursor Analogs
29 FILE REFERENCE: 6058,200-US
30 CURRENT APPLICATION NUMBER: US/09/736,611
31 CURRENT FILING DATE: 2000-12-14
32 PRIOR APPLICATION NUMBER: 60/181,443
33 PRIOR FILING DATE: 2000-02-10
34 PRIOR APPLICATION NUMBER: 60/211,441
35 PRIOR FILING DATE: 2000-06-13
36 PRIOR APPLICATION NUMBER: PA 1999 01888
37 PRIOR FILING DATE: 1999-12-29
38 PRIOR APPLICATION NUMBER: PA 2000 00440
39 PRIOR FILING DATE: 2000-03-17
40 NUMBER OF SEQ ID NOS: 18
41 SOFTWARE: FastSeq for Windows Version 4.0
42 SEQ ID NO 10
43 LENGTH: 125
44 TYPE: PRT
45 ORGANISM: N-terminal extension
46
47 US-09-736-611-10
48
49 Query Match 46.9%; Score 275.5; DB 9; Length 125;
50 Best Local Similarity 90.9%; Pred. No. 2,1e-25;
51 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
52
53
54 54 PFVNOHLCGSHLVALYIVCGERGFFVTPKT-RGIVEQCCTSIICSLYOLENYCN 107
55 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 71 PFVNOHLCGSHLVALYIVCGERGFFVTDKKGIVEQCCTSIICSLYOLENYCN 125
57
58
59 RESULT 14
60 US-09-740-353-10
61 Sequence 10, Application US/09740359
62 Patient No. US2C010041787A1
63 GENERAL INFORMATION:
64 APPLICANT: Kjedtsen, Thomas Borglum
65 APPLICANT: Ludvigsen, Svend
66 TITLE OF INVENTION: Method for making insulin precursors and
67 TITLE OF INVENTION: Insulin precursor analogues having improved fermentation
68 TITLE OF INVENTION: yield in yeast
69 FILE REFERENCE: 6148,200-US
70 CURRENT APPLICATION NUMBER: US/09/740,359
71 CURRENT FILING DATE: 2000-12-19
72 PRIOR APPLICATION NUMBER: PA 2000 00443
73 PRIOR FILING DATE: 2000-03-17
74 PRIOR APPLICATION NUMBER: PA 1999 01889
75 PRIOR FILING DATE: 1999-12-29
76 PRIOR APPLICATION NUMBER: 60/211,081
77 PRIOR FILING DATE: 2000-06-13

```

```

1 LENGTH: 124
2 TYPE: PRT
3 ORGANISM: Artificial Sequence
4 FEATURE:
5 OTHER INFORMATION: Synthetic
6
7 US-09-894-711-12
8
9 Query Match 46.9%; Score 275.5; DB 10; Length 124;
10 Best Local Similarity 90.9%; Pred. No. 2.1e-25;
11 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
12
13
14 54 PFVNOHLCGSHLVALYIVCGERGFFVTPKT-RGIVEQCCTSIICSLYOLENYCN 107
15 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 70 PFVNOHLCGSHLVALYIVCGERGFFVTDKKGIVEQCCTSIICSLYOLENYCN 124
17
18
19 RESULT 13
20 US-09-740-611-10
21 Sequence 10, Application US/09736611
22 Patient No. US2C010023069A1
23 GENERAL INFORMATION:
24 APPLICANT: Kjedtsen, Thomas
25 APPLICANT: Ludvigsen, Svend
26 APPLICANT: Kaarshoim, Niels
27 TITLE OF INVENTION: Method For Making Insulin Precursors and
28 TITLE OF INVENTION: Insulin Precursor Analogs
29 FILE REFERENCE: 6058.200-US
30 CURRENT APPLICATION NUMBER: US/09/736,611
31 CURRENT FILING DATE: 2000-12-14
32 PRIOR APPLICATION NUMBER: 60/181,443
33 PRIOR FILING DATE: 2000-02-10
34 PRIOR APPLICATION NUMBER: 60/211,441
35 PRIOR FILING DATE: 2000-06-13
36 PRIOR APPLICATION NUMBER: PA 1999 01888
37 PRIOR FILING DATE: 1999-12-29
38 PRIOR APPLICATION NUMBER: PA 2000 00440
39 PRIOR FILING DATE: 2000-03-17
40 NUMBER OF SEQ ID NOS: 18
41 SOFTWARE: FastSeq for Windows Version 4.0
42 SEQ ID NO 10
43 LENGTH: 125
44 TYPE: PRT
45 ORGANISM: N-terminal extension
46
47 US-09-736-611-10
48
49 Query Match 46.9%; Score 275.5; DB 9; Length 125;
50 Best Local Similarity 90.9%; Pred. No. 2.1e-25;
51 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
52
53
54 54 PFVNOHLCGSHLVALYIVCGERGFFVTPKT-RGIVEQCCTSIICSLYOLENYCN 107
55 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 71 PFVNOHLCGSHLVALYIVCGERGFFVTDKKGIVEQCCTSIICSLYOLENYCN 125
57
58
59 RESULT 14
60 US-09-740-353-10
61 Sequence 10, Application US/09740359
62 Patient No. US2C010041787A1
63 GENERAL INFORMATION:
64 APPLICANT: Kjedtsen, Thomas Borglum
65 APPLICANT: Ludvigsen, Svend
66 TITLE OF INVENTION: Method for making insulin precursors and
67 TITLE OF INVENTION: Insulin precursor analogues having improved fermentation
68 TITLE OF INVENTION: yield in yeast
69 FILE REFERENCE: 6148.200-US
70 CURRENT APPLICATION NUMBER: US/09/740,359
71 CURRENT FILING DATE: 2000-12-19
72 PRIOR APPLICATION NUMBER: PA 2000 00443
73 PRIOR FILING DATE: 2000-03-17
74 PRIOR APPLICATION NUMBER: PA 1999 01869
75 PRIOR FILING DATE: 1999-12-29
76 PRIOR APPLICATION NUMBER: 60/211,081
77 PRIOR FILING DATE: 2000-06-13

```



```
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 125
; TYPE: PRI
; ORGANISM: TA39 leader fused with N-terminally extended
US-09-740-359-10

Query Match          46.9%  Score 275.5; DB 9; Length 125;
Best Local Similarity 90.9%  Pred. No. 2.1e-25;
Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 54 PRFVNQHLGSHLVEALYLVCGERGFFYTPKT-RGIVEGCCISCSLYQLNRYCN 107
    I:|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I||
Db 71 PRFVNQHLGSHLVEALYLVCGERGFFYTDKDKGIVEGCCISCSLYQLNRYCN 125
    I:|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I||

RESULT 15
US-09-894-711-10
; Sequence 10, Application US/09894711
; Patent No. US2002037144A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Reqlue
; APPLICANT: Jadvigsen, Svend
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursor analogues having improved fortification
; TITLE OF INVENTION: yield in yeast
; FILE REFERENCE: 6148,400-US
; CURRENT APPLICATION NUMBER: US/09/894,711
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01859
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 09/740,359
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 125
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-894-711-10
```

```
Query Match          46.9%  Score 275.5; DB 10; Length 125;
Best Local Similarity 90.9%  Pred. No. 2.1e-25;
Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 54 PRFVNQHLGSHLVEALYLVCGERGFFYTPKT-RGIVEGCCISCSLYQLNRYCN 107
    I:|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I||
Db 71 PRFVNQHLGSHLVEALYLVCGERGFFYTDKDKGIVEGCCISCSLYQLNRYCN 125
    I:|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I||
```

Search completed: September 16, 2003, 12:52:26
Job time : 34.7237 secs

GeneCore version 5.1.5
Copyright (c) 1993 - 2003 Computer Ltd.

OK protein - protein search, using sw model

Run on: September 16, 2003, 12:36:15 Search time 12.4603 seconds
(without alignments)
823,645 million cell updates/sec

Title: US-09-423-100-6

Perfect score: 597

Sequence: 1 MEPIPIPSRLFDNAVLRAHR.....IVRQVTSIGSLYLENYCN 107

Scoring table: BL0SUM62

Gapop 10.0 Gapext 0.5

Searched: 283308 seqs, 96166882 residues

Total number of hits satisfying chosen parameters: 261828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR_76:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	275	46.8	96	2	PC7082
2	273.5	46.6	51	1	INRHF
3	273.5	46.6	51	1	INRHF
4	273.5	46.6	51	1	INRHF
5	273	45.5	110	2	B42179
6	273	46.5	110	2	QC178
7	271.5	46.3	51	1	INRHF
8	268.5	45.7	51	1	INRHF
9	267.5	45.6	51	2	A9131
10	267	45.5	110	1	INRHF
11	267	45.5	110	2	A42179
12	263.5	44.9	51	1	INRHF
13	263.5	44.9	51	1	INRHF
14	263.5	44.9	51	1	INRHF
15	263	44.8	84	1	INRHF
16	263	44.8	84	1	INRHF
17	262.5	44.7	110	1	INRHF
18	262	44.6	110	1	INRHF
19	261.5	44.5	51	1	INRHF
20	260	44.3	110	2	I48156
21	258.5	44.0	105	1	INRHF
22	257	43.8	108	2	A93883
23	256.5	43.7	51	2	C03362
24	255.5	43.5	217	1	STHC
25	255.5	43.5	217	2	I57410
26	252.5	43.0	77	1	INRHF
27	252	42.9	86	1	INRHF
28	251.5	42.8	51	1	INRHF
29	250	42.6	108	1	INRHF

```

30      249      42.4      110      1      IPRI1      insulin 1 precursor
31      248.5      42.3      51      1      INGS      insulin - goose
32      248      42.2      110      1      IPRI2      insulin 2 precursor
33      248      42.2      110      1      INMS2      insulin 2 precursor
34      246      41.9      52      2      S44469      insulin 11 - North
35      246      41.9      52      2      S44470      insulin 12 - North
36      245      41.7      103      2      I51221      insulin precursor
37      244.5      41.7      51      1      INPO      insulin - crested
38      244.5      41.7      51      1      INTK      insulin - turkey
39      244.5      41.7      51      1      INOS      insulin - ostrich
40      244.5      41.7      51      1      A61129      insulin - black-be
41      244.5      41.7      51      2      A60414      insulin - slider t
42      239.5      40.8      137      1      IPCH      insulin precursor
43      238      40.5      52      2      S61361      insulin - Amphiuma
44      235.5      40.1      51      2      S63590      insulin - duckbill
45      233.5      39.8      91      1      IPDK      insulin precursor

```

ALIGNMENTS

RESULT 1

PC7082

Epidermal growth factor/single chain insulin fusion protein - Bacillus brevis (fragr
Accession: PC7082; sequence_revision 18-Aug-2000 #text_change 31-Mar-2003
Accession: PC7082; PC7083
Hiroh, H.; Haragata, H.; Ebisu, S.; Morihara, K.; Takagi, H.
Biosci. Biotechnol. Biochem. 64, 1079-1081, 2000
A>Title: Use of Bacillus brevis for synthesis and secretion of Des-B30 single-chain
AReference number: PC7082; NCBI:20335834; PMID:10879487
AAccession: PC7082
A.Molecule type: DNA
A.Residues: 1-96 <KOH>
A.Accession: PC7083
A.Molecule type: protein
A.Residues: 19-26 <K02>
C:Genetics:
A:Gene: eaf-sci
C:Superfamily: insulin

Query Match 46.8% Score 275; DB 2; Length: 96;

Best Local Similarity 94.3%; Pred. No. 1.2e-21;

Matches 50; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 55 KFYNOHLGSGHVEALYLVGGERGFEYTKTRIGVGCCTICSLYLENYCN 107

DB 46 KFYNOHLGSGHVEALYLVGGERGFEYTKR--GLVEGCTICSLYLENYCN 96

RESULT 2

INRHF

insulin - sperm whale

C:Species: Physocera catodon (sperm whale)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #ext_change 16-Jun-1999

C:Accession: A93142; A90082

Kishikawa, Y.; Saito, T.; Ito, Y.; Fujino, M.

Nature 361, 1468-1469, 1998

A>Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pef

AReference number: A93142

AAccession: A93142

A.Molecule type: protein

A.Residues: 1-30; 31-51 <ISH>

P.Harris, J.L.; Sanger, F.; Naughton, M.A.

Arch. Biochem. Biophys. 55, 427-428, 1956

A>Title: Species differences in insulin.

AReference number: A90082

AAccession: A90082

A.Molecule type: protein

A.Residues: 1-30; 31-51 <HAR>

C:Superfamily: insulin

C:Keywords: hormone; pancreas

F/1-30/Domain: Insulin chain B #status experimental <BCH>

J. Biol. Chem. 246, 1375-1386, 1971
A:Title: Studies on human proinsulin: isolation and amino acid sequence of the human pan-
A:Reference number: A92075; MUID:71116450; PMID:5101771
A:Accession: A92075
A:Molecule type: protein
A:Residues: 57-87 <OE>
R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 190-199, 1971
A:Title: Amino acid sequence of the C-peptide of human proinsulin.
A:Reference number: A91186; MUID:7157722; PMID:5560424
A:Accession: A91186
A:Molecule type: protein
A:Residues: 57-87 <OA>
R:Lucassen, A.M.; Giller, C.; Heressi, J.P.; Roillard, G.; Frossel, P.; Rathrop, M.; Bell
Nature Genet. 4, 305-310, 1993
A:Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment.
A:Reference number: I58114; MUID:93364428; PMID:8458440
A:Accession: I58114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-59,63-110 <RFS>
A:Cross-references: GB:115440; NID:9307371; PID:AAA59179.1; PID:9307372
R:Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Riniker, B.; Rittel, W.
Helv. Chim. Acta 57, 2617-2621, 1974
A:Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
A:Reference number: A91636; MUID:75077277; PMID:4443293
A:Contents: annotation; synthesis
A:Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was found
R:Naithani, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A:Title: The synthesis of C-peptide of human proinsulin.
A:Reference number: A91658; MUID:75040007; PMID:4404564
A:Contents: annotation; synthesis of residues 57-87
R:Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 105, 2347-2352, 1973
A:Title: Synthesis of the complete sequence of human proinsulin C-peptide and its C-
A:Reference number: A90914
A:Contents: annotation; synthesis of residues 57-87
R:Kaufmann, J.E.; Iminger, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A:Title: Sequence requirements for proinsulin processing at the P-chain/peptide junction
A:Reference number: S58661; MUID:96013185; PMID:7575420
A:Contents: annotation; site-directed mutagenesis study of propeptidic processing
C:Genetics:
A:Gene: GDR:INS
A:Cross-references: GDR:119343; OMIM:176730
A:Map position: 11p15.5 11p15.5
A:Features: 63/
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status experimental <SCH>
F:25-54/Domain: insulin chain B #status experimental <SCH>
F:55-84,90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting C-peptide #status experimental <CPPT>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status experimental
Query Match 45.5%; Score 267; DB 1; Length 110;
Best Local Similarity 60.5%; Pred. No. 9,3e-21;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
QY 56 FVNHLCGSHIVEALYLVCGGFFYTPKT----- 85
DB 25 FVNHLCGSHIVEALYLVCGGFFYTPKTREARDLQGVGLGGSPGAGSLQPLALEG 84
QY 86 ----RGIVFOCCSTCSLYOLENYCN 107
DB 85 SLOKRGIVECCSTCSLYOLENYCN 110
RESULT 12
insulin - sei whale
C:Species: Balanoptera borealis (sei whale)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C:Accession: A01582
R:Shihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 311, 1458-1469, 1988
A:Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pe
A:Reference number: A93142
A:Accession: A01582
A:Molecule type: protein
A:Residues: 1-30;31-51 <ISH>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <ACH>
F:31-51/Product: insulin #status experimental <MAT>
F:51-51/Domain: insulin chain A #status experimental <ACH>
F:37-37,19-50,35-41/Disulfide bonds: #status predicted
Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 9.6e-21;
Matches 49; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 56 FVNHLCGSHIVEALYLVCGGFFYTPKTREARDLQGVGLGGSPGAGSLQPLALEG 107
DB 1 FVNHLCGSHIVEALYLVCGGFFYTPKA-GIVECCASTCSLYOLENYCN 51
RESULT 13
insulin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C:Accession: A01586
R:Smith, L.P.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-30;31-51 <SM1>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model:

Run on: September 16, 2003, 12:33:30 : Search time 8.2285 Seconds
(without alignments)
604,291 Million cell updates/sec

Title: US-09-423-100-6

Percent score: 567

Sequence:

1 MFPIFLSLPUNPMLRAHP.....IVEGQGISLCELYAEVENVN 107

Scoring table:

EJOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127853 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	273.5	46.6	51	INS_BALPH	P01312 balaenopter
2	273.5	46.6	51	INS_FELCA	P01316 elephas max
3	273	46.5	110	INS_CERAE	P30407 ceropithec
4	273	46.5	110	INS_MACFA	P30406 macaca fasc
5	268.5	45.7	51	INS_ACCOA	P01324 acorys cabi
6	267	45.5	110	INS_HUMAN	P01368 homo sapien
7	267	45.5	110	INS_PANTR	P30419 pan troglod
8	266	45.3	110	INS_SPEPR	P01313 spermophilu
9	263.5	44.9	51	INS_FALBO	P01314 balaenopter
10	263.5	44.9	51	INS_CABOR	P01320 camelus dro
11	263.5	44.9	51	INS_CABHI	P01319 capra airo
12	263	44.8	108	INS_FIG	P01315 sus scrofa
13	263	44.8	110	INS_RABIT	P01311 corytophila
14	262.5	44.7	51	INS_FELCA	P06306 felis stiro
15	262	44.6	110	INS_CANFA	P01321 canis fami
16	260	44.3	110	INS_CRILG	P01317 crinophila
17	258.5	44.0	105	INS_BOVIN	P01317 los tauru
18	257	43.8	108	INS_AOTIR	P00604 aotus triu
19	257	43.8	110	INS_PSAOB	P02587 psammomy
20	256.5	43.7	51	INS_DIDKA	P01309 didelphis m
21	255.5	43.5	217	SOMA_HUMAN	P02541 homo sapien
22	255.5	43.5	217	SOMA_MACMO	P33333 macaca mul
23	255.5	43.5	217	SOMA_PANTR	P58756 pan troglod
24	254.5	43.4	105	INS_SREBP	P01318 ovis arie
25	253	42.9	86	INS_HORSE	P01310 equus cabi
26	251.5	42.8	51	INS_CHARR	P01327 chinchilla
27	250	42.6	108	INS_MOUSE	P01325 mus muscu
28	249	42.4	110	INS1_RAT	P01322 ratu
29	249	42.4	217	SOMA_CALJA	P08411 callositrix
30	249	42.4	217	SOMA_SALB3	P58411 salicaria
31	248.5	42.3	51	INS_ANSAN	P03454 anser anser
32	248	42.2	110	INS2_MOUSE	P01323 mus muscu
33	248	42.2	110	INS2_RAT	P01323 ratu

34 246 41.9 52 1 INS_ACIQU
35 245 41.7 103 1 INS_SELRF
36 244.5 41.7 51 1 INS_HYSCR
37 244.5 41.7 51 1 INS_TRASC
38 239.5 40.8 107 1 INS_CHICK
39 236 40.2 217 1 SOM2_PANTR
40 235.5 39.8 51 1 INS_ORNAN
41 233.5 39.8 81 1 INS_ANAPL
42 231.5 39.4 51 1 INS_ALUMI
43 231 39.4 52 1 INS_LEPSP
44 228.5 38.9 51 1 INS_ZAODH
45 228 38.8 217 1 SOM2_HUMAN

P01423 acipenser g
P01463 selasphorus
P01328 hystrix cri
P01887 trachemys s
P01332 gallus gall
P08757 pan troglod
P09497 ornithotyr
P01333 anas platyr
P02703 alligator m
P09476 lepisosteus
P01208 zaocys dhum
P01242 homo sapien

ALIGNMENTS

RESULT 1
INS_BALPH STANDARD: PRT: 51 AA.
AC P01312
CT 21-JUL-1986 (Ref. 31, Created)
CT 21-JUL-1986 (Ref. 31, Last sequence update)
CT 21-JUL-1986 (Ref. 34, Last annotation update)
34 Insulin.
35 INS.
36 Balaenoptera physalus (Finback whale) (Common rorqual), and
37 Physeter catodon (Sperm whale) (Physeter macrocephalus).
38 Eukaryota: Metazoa: Chordata: Cranialia: Vertebrata: Euteleostomi:
39 Mammalia: Cetartiodactyla: Cetacea: Mysticeti:
40 Balaenopteridae: Balaenoptera.
41 NCBI_TaxID=9770, 9755;
42
43 PARTIAL SEQUENCE.
44 SPECIES=B.physalus;
45 KA Rama R., Titan K., Sakaki S., Norita K.;
46 "The amino acid sequence in fin-whale insulin.";
47 J. Biochem. 56:285-293(1964).
48 (2)
49 SEQUENCE.
50 SPECIES=P.catodon;
51 Ishihara Y., Saito T., Ito Y., Fujino M.;
52 "Structure of sperm- and sei-whale insulins and their breakdown by
53 whale pepsin.";
54 Nature 161:1468-1469(1958).
55
56 SEQUENCE.
57 SPECIES=P.catodon;
58 Harris J.I., Sander P., Naughton M.A.;
59 "Species differences in insulin.";
60 Arch. Biochem. Biophys. 65:427-438(1956).
61
62 FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
63 INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
64 FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
65 CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
66 SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
67 DISULFIDE BONDS.
68
69 -1- SUBCELLULAR LOCATION: Secreted.
70
71 -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
72 PIR: A91918; INWHF.
73 PIR: A91142; INWHF.
74 HSSP: P01317; IAPB.
75 InterPro: IPR004325; Ins/IGF/relax.
76 SMART: SM00078; IIGF. 1
77 PROSITE: PS00262; INSULIN: 1.
78 Insulin family; Hormone; Glucose metabolism.
79 CHAIN 1 30
80 NON_CONS 30 31
81 CHAIN 31 51
82 INSULIN A CHAIN.
83 DISULFID 7 37
84 INTERCHAIN.
85 DISULFID 13 50
86 DISULFID 36 41
87 SEQUENCE 51 AA: 5766 MW: 9007514691A7CDD CRC64;

```

Query Match      46.64; Score 273.5; DS 1; Length 110;
Best Local Similarity 96.24; Pred. No. 166-22;
Matches 50; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 107
DB 1 FVNQHLGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 82

RESULT 2
INS_ELEMA
ID INS_ELEMA STANDARD: PRT: 51 AA.
AC P01316;
DT 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin.
GN INS.
OS Elephas maximus (Indian elephant).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
OX NCBI_TaxID=9783;
RN [1]
RP SEQUENCE.
RX MEDLINE=66160119; PubMed=5049593;
RA Smith L.F.;
RL Am. J. Med. 40:662-666(1966);
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X67092; CAA43405.1;
CC FTK: B42179; B42179.
CC HSP: P01308; IALC.
CC InterPro: IPR004825; Ins/IGF/relax.
CC SMART: PF000649; Insulin; 1.
CC SMART: SMC0078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family; Hormone; Glucose metabolism; Signal.
CC CHAIN 1 24 INSULIN B CHAIN.
CC NON_CONS 25 54 C PEPTIDE.
CC PROPEP 57 87
CC CHAIN 90 110 INSULIN A CHAIN.
CC DISULFID 31 96 INTERCHAIN.
CC DISULFID 43 104 INTERCHAIN.
CC DISULFID 95 100
CC SEQUENCE 51 AA; 5752 MW; 900755008457100 CRC64;

Query Match      46.64; Score 273.5; DS 1; Length 110;
Best Local Similarity 96.24; Pred. No. 166-22;
Matches 50; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 107
DB 1 FVNQHLGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 82

RESULT 3
INS_CERAE
ID INS_CERAE STANDARD: PRT: 110 AA.
AC P30407; P01309;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

```

```

Cercopithecoidea; Cercopithecoidea.
NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219953; PubMed 1560757;
FA Seino S., Bell G.L., Li W.;
FT "Sequences of primate insulin genes support the hypothesis of a
FT slower rate of molecular evolution in humans and apes than in
FT monkeys.";
FL Mol. Biol. Evol. 9:193-203(1992).
RN [2]
RP SEQUENCE OF 57-87.
RX MEDLINE=72258016; PubMed=4626369;
FA Peterson C.D., Nehrlich S., Czer P.E., Steiner D.F.;
FT "Determination of the amino acid sequence of the monkey, sheep, and
FT dog proinsulin C-peptides by a semi-micro Edman degradation
FT procedure.";
FL J. Biol. Chem. 247:4866-4871(1972).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X67092; CAA43405.1;
CC FTK: B42179; B42179.
CC HSP: P01308; IALC.
CC InterPro: IPR004825; Ins/IGF/relax.
CC SMART: PF000649; Insulin; 1.
CC SMART: SMC0078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family; Hormone; Glucose metabolism; Signal.
CC CHAIN 1 24 INSULIN B CHAIN.
CC NON_CONS 25 54 C PEPTIDE.
CC PROPEP 57 87
CC CHAIN 90 110 INSULIN A CHAIN.
CC DISULFID 31 96 INTERCHAIN.
CC DISULFID 43 104 INTERCHAIN.
CC DISULFID 95 100
CC SEQUENCE 110 AA; 12319 MW; 95A1F54BE7B24749 CRC64;

Query Match      46.54; Score 273; DS 1; Length 110;
Best Local Similarity 96.24; Pred. No. 41e-22;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 FPNVNHGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 85
DB 23 FPNVNHGGSHLVEALYLVCGERGFFYTPKTRGVGQVQLVQDGLPGQVQSIVQLPQLR 82

QY 86 -----RGIVGQVQLVQDGLPGQVQSIVQLPQLR 107
DB 83 EGSQKRGIVGQVQLVQDGLPGQVQSIVQLPQLR 110

RESULT 4
INS_MACFA
ID INS_MACFA STANDARD: PRT: 110 AA.
AC P30406; P01309;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.

```



```

GN  INS.
OS  Macaca fascicularis (Orang eating macaque) (Cynomolgus monkey).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Ceropithecidae;
OC  Cercopithecinae; Macaca.
OX  NCBI_TaxID=9547;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=83060474; PubMed=6154262;
RA  Weinreb W., Groneberg J., Weinreb M., Weinreb P.,
RA  Winacker E.;
RT  "The nucleotide sequence of cDNA coding for preproinsulin from the
RL  Gene 19,179-183(1982).
CC  -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC  INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC  FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC  CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC  -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC  DISULFIDE BONDS.
CC  -!- SURCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is not
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/affiliates/
CC  or send an email to license@sib.ch).
CC  -----
DR  EMBL: J00135; AAA16849.1;
DR  PIR: J00175; J00178;
DR  HSSP: P01308; TAIG.
DR  InterPro: IPR004925; Ins/IGF/relax.
DR  Print: P00049; Insulin; 1.
DR  PRANTS: PR0277; INSULINS.
DR  SMART: SM0076; IGF; 1.
DR  PROSITE: PS00262; INSULIN; 1.
KW  Insulin family; Hormone; Glucose metabolism; Signal.
FT  SIGNAL 1 24
FT  CHAIN 25 54 INSULIN B CHAIN.
FT  PROPEP 57 57 C PEPTIDE.
FT  CHAIN 90 110 INSULIN A CHAIN.
FT  DISULFID 31 95 INTERCHAIN.
FT  DISULFID 43 103 INTERCHAIN.
FT  DISULFID 95 100
FT  SEQUENCE 110 AA; 11991 MW; 8306E33A80A420F9 CRC64;
Query Match 46.5%; Score 273; DB 1; Length 110;
Best Local Similarity 60.2%; Pred. No. 4,1e-22;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PAFVNHGCGSHLVLEALYVCGERGFFYTPKTR-----
DB 23 PAFVNHGCGSHLVLEALYVCGERGFFYTPKTRFAEDFVQVQVVGGSFAGNSQLEAL 22
QY P5 -----RGVVGQCTGCSLYQLENVCN 107
DB 63 EASLQKRGVVGQCTGCSLYQLENVCN 110
RESULT 5
ID INS_ACOCA STANDARD; 1RT; 51 AA.
AC P01324;
DT 21-JUL-1986 (Rel. 01; Created)
DT 21-JUL-1986 (Rel. 01; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE Insulin.
GN INS.
OS Acomys cahirinus (Egyptian spiny mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Acomys.
OX  NCBI_TaxID=10068;
RN  [1]
RP  COMPOSITION.
RX  MEDLINE=72189454; PubMed=5028210;
RA  Rucznik H.F., Hummel R.E.;
RA  "Isolation and partial structural analysis of insulin from mouse (Mus
RA  musculus) and spiny mouse (Acomys cahirinus).";
RL  Hoppe-Seyler's Z. Physiol. Chem. 353:444-450(1972).
CC  -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC  INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC  FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC  CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC  -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC  DISULFIDE BONDS.
CC  -!- SURCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR  PIR: AC1591; INHSSP.
DR  HSSP: P01308; ITYM.
DR  InterPro: IPR004925; Ins/IGF/relax.
DR  SMART: SM0076; IGF; 1.
DR  PROSITE: PS00262; INSULIN; 1.
KW  Insulin family; Hormone; Glucose metabolism.
FT  CHAIN 30 50 INSULIN B CHAIN.
FT  NOKCONS 30 50
FT  CHAIN 31 51 INSULIN A CHAIN.
FT  DISULFID 37 51 INTERCHAIN (BY SIMILARITY).
FT  DISULFID 49 50 INTERCHAIN (BY SIMILARITY).
FT  DISULFID 36 41 BY SIMILARITY.
FT  SEQUENCE 51 AA; 5768 MW; 992ED8B6290473DB CRC64;
Query Match 45.7%; Score 262.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 5,5e-22;
Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNHGCGSHLVLEALYVCGERGFFYTPKTRGIVEGQCTGCSLYQLENVCN 107
DB 111 FVNHGCGSHLVLEALYVCGERGFFYTPKTRGIVEGQCTGCSLYQLENVCN 111
DB 1 FVNHGCGSHLVLEALYVCGERGFFYTPKS-GIVDQCTGCSLYQLENVCN 51
RESULT 6
ID INS_HUMAN STANDARD; PRT; 110 AA.
AC P01308;
DT 21-JUL-1986 (Rel. 01; Created)
DT 21-JUL-1986 (Rel. 01; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Insulin precursor.
GN INS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=80120725; PubMed=6243748;
RA  Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.;
RA  Goodman H.M.;
RT  "Sequence of the human insulin gene.";
DB  Nature 284:26-32(1980).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=80236313; PubMed=6248962;
RA  Ullrich A., Dull T.C., Gray A., Brosius J., Sures I.;
RA  "Genetic variation in the human insulin gene.";
DB  Science 239:612-615(1980).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=80054775; PubMed=501234;
RA  Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
RA  Rutter W.J.;
RT  "Nucleotide sequence of a cDNA clone encoding human preproinsulin.";
DB  Nature 282:525-527(1979).

```

RN [4] SEQUENCE FROM N.A.
 RP MEDLINE=8214717; PubMed=6927840;
 RA Sures I., Goeddel D.V., Gray A., Ullrich A.;
 RT "Nucleotide sequence of human proinsulin complementary DNA";
 RL Science 238:57-59(1989).
 RN [5] SEQUENCE FROM N.A.
 RP MEDLINE=93364428; PubMed=358640;
 RA Lucassen A.M., Bell J.I., Joller G., Lathrop M.;
 RT "Susceptibility to insulin dependent diabetes mellitus maps to a 4 kb segment of DNA spanning the insulin gene and associated VNTR";
 RL Nat. Genet. 4:305-310(1994).
 RN [6] SEQUENCE FROM N.A.
 RP TISSUE=Pancreas;
 RC MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derce J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Euetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hooq B., Stapleton M., Soares M.B., Donaldo M.F., Casavant T.L., Scheetz T.E., Raba S.S., Loqueilaco N.A., Peters G.J., Abranson E.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Marek J.A., Gumarane P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay S.J., Hulyx S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton A., Kottmann M., Madan A., Rodriguez S., Sanchez A., Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bonfield W.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywicki M.I., Skalska H., Smalius D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7] SEQUENCE OF 1-59 FROM N.A.
 RP TISSUE=Blood;
 RC MEDLINE=71113410; PubMed=5161771;
 RA Fajardy I.L., Weil D.J., Stackens C.G., Panse E.M.P.;
 RT "Description of a novel KRP diallelic polymorphism (-127 T>G) within the 5' region of insulin gene";
 RL Submitted (JUL-1998) to the ENR/GenBank/EMBL databases.
 RN [8] SEQUENCE OF 25-54 AND 90-110.
 RP MEDLINE=7125722; PubMed=5560404;
 RA Nicoi D.S.H.W., Smith J.F.;
 RT "Amino-acid sequence of human insulin";
 RL Nature 187:483-485(1960).
 RN [9] SEQUENCE OF 57-87.
 RP MEDLINE=71113410; PubMed=5161771;
 RA Oyer P.E., Cho S., Peterson J.D., Steiner D.F.;
 RT "Studies on human proinsulin. Isolation and amino acid sequence of the human pancreatic C-peptide";
 RL J. Biol. Chem. 246:1375-1383(1971).
 RN [10] SEQUENCE OF 57-87.
 RP MEDLINE=7125722; PubMed=5560404;
 RA Ko A., Smyth D.G., Markusson J., Sundby F.;
 RT "The amino acid sequence of the C-peptide of human proinsulin";
 RL Eur. J. Biochem. 20:190-193(1971).
 RN [11] SEQUENCE OF 57-87.
 RP MEDLINE=75077277; PubMed=4443293;
 RA Sieber P., Kamber B., Hartmann A., Geor A., Zinkler B., Kitterl W.;
 RT "Total synthesis of human insulin order directed formation of the disulfide bonds";
 RL Helv. Chim. Acta 57:2617-2621(1974).
 RN [12] SEQUENCE OF 57-87.
 RP MEDLINE=75040007; PubMed=48(3504);
 RA Nathani V.K.;
 RT "Studies on polypeptides. IV. The synthesis of C-peptide of human proinsulin";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
 RN [13] SYNTHESIS OF 65-69 AND 70-73.
 RP MEDLINE=73161283; PubMed=4598555;
 RA Geiger R., Volk A.;
 RT "Synthesis of peptides with the properties of human proinsulin C peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 9-13 of human proinsulin C peptides";
 RL Chem. Ber. 106:199-205(1973).
 RN [14] SYNTHESIS OF 84-87.
 RP MEDLINE=73161261; PubMed=4598553;
 RA Geiger R., Jaeger G., Keonig W., Treuth G.;
 RT "Synthesis of peptides with the properties of human proinsulin C peptides (hC peptide). 1. Scheme for the synthesis and preparation of the sequence 28-31 of human proinsulin C peptide";
 RL Chem. Ber. 106:188-192(1973).
 RN [15] VARIANT LOS ANGELES SER-48.
 RP MEDLINE=84016053; PubMed=6312455;
 RA Hameda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
 RT "Studies on mutant human insulin genes: identification and sequence analysis of a gene encoding [SerB24]insulin";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
 RN [16] VARIANTS LOS ANGELES SER-48 AND CHICAGO LEU-49.
 RP MEDLINE=84170233; PubMed=6424111;
 RA Shoelson S., Fickova M., Hameda M., Nahum A., Musso G., Kaiser E.T., Rubenstein A.H., Tager H.;
 RT "Identification of a mutant human insulin predicted to contain a serine-for-phenylalanine substitution";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
 RN [17] VARIANT PROVIDENCE ASP-34.
 RP MEDLINE=87175640; PubMed=3470784;
 RA Chan S.J., Seino S., Gruppiso P.A., Schwartz K., Steiner D.F.;
 RT "A mutation in the B chain coding region is associated with impaired proinsulin conversion in a family with hyperproinsulinemia";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
 RN [18] VARIANT WAKAYAMA LEU 92.
 RP MEDLINE=87056122; PubMed=3537011;
 RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
 RT "Structurally abnormal insulin in a diabetic patient. Characterization of the mutant insulin A3 (Val-->Leu) isolated from the pancreas";
 RL J. Clin. Invest. 78:1666-1672(1986).
 RN [19] VARIANT HIS-89.
 RP MEDLINE=90317021; PubMed=2196279;
 RA Barbelli F., Raben N., Kadowaki T., Cama A., Accilli D., Gabbay K.H., Merenich J.A., Taylor S.L., Roth J.;
 RT "Two unrelated patients with familial hyperproinsulinemia due to a mutation substituting histidine for arginine at position 65 in the proinsulin molecule: identification of the mutation by direct sequencing of genomic deoxyribonucleic acid amplified by polymerase chain reaction";
 RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
 RN [20] VARIANT HIS-89.
 RP MEDLINE=85261946; PubMed=4019786;
 RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
 RT "Posttranslational cleavage of proinsulin is blocked by a point mutation in familial hyperproinsulinemia";
 RL J. Clin. Invest. 76:378-380(1985).
 RN [21] VARIANT KYOTO LEU-89.
 RP MEDLINE=92291307; PubMed=1601997;
 RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Izura H., Seino Y.;
 RT "A novel point mutation in the human insulin gene giving rise to hyperproinsulinemia (proinsulin Kyoto)";
 RL J. Clin. Invest. 89:1902-1907(1992).

```

RN [22]
RP STRUCTURE BY NMR.
RX MEDLINE-91104966; PubMed-2271664;
RA Hua Q.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RT structure."
RL Biochemistry 29:10545-10555(1990).
RN [23]
RP STRUCTURE BY NMR.
RX MEDLINE-91242457; PubMed-2036426;
RA Hua Q.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition."
RL Biochemistry 30:5505-5515(1991).
RN [24]
RP STRUCTURE BY NMR.
RX MEDLINE-91265527; PubMed-1846635;
RA Hua Q.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Des-(526-530)-insulin: sequence-
RT specific resonance assignments and effects of solvent composition."
RL Biochim. Biophys. Acta 1078:101-110(1991).

Query Match 45.5%; Score 267; DB 1; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.8e-21;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNHGCGSHVEALYLVCGERGFFYTPKT----- 85
DB 25 FVNHGCGSHVEALYLVCGERGFFYTPKTRERAEHLQVGVGLGAGSGIOPLALES 84
QY 86 ----RGIVEQCTSIQSLQVLENYCN 107
DB 85 SLCKRGIVEQCTSIQSLQVLENYCN 110

RESULT 6
INS_PANTR
ID INS_PANTR STANDARD; PRT; 110 AA.
P30410.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCB:_TaxID=9598;
RN [1]
RX MEDLINE-92219553; PubMed-1560757;
RA Saino S., Bell G.T., Li W.;
RT "Sequences of primate insulin genes support the hypothesis of a
RT slower rate of molecular evolution in humans and apes than in
RT monkeys."
RL Mol. Biol. Evol. 9:193-203(1992).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
CC
CC EMBL: X61089; CAA43403.1;
CC
CC PIR: A42179; A42179.
CC
CC InterPro: IPR004825; Ins/IGF/relax.
CC
CC PRINTS: PR00277; INSULINB.
CC
CC SMART: SM00078; IGF; 1.
CC
CC PROSITE: PS00252; INSULIN; 1.
CC
CC Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12025 MW; 41EB8DF79837CEP5 CRC64;

Query Match 45.5%; Score 267; DB 1; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.8e-21;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNHGCGSHVEALYLVCGERGFFYTPKT----- 85
DB 25 FVNHGCGSHVEALYLVCGERGFFYTPKTRERAEHLQVGVGLGAGSGIOPLALES 84
QY 86 ----RGIVEQCTSIQSLQVLENYCN 107
DB 85 SLCKRGIVEQCTSIQSLQVLENYCN 110

RESULT 6
INS_PSETR
ID INS_PSETR STANDARD; PRT; 110 AA.
Q91X13.
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Spornophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spermophilus.
CC NCB:_TaxID=43179;
RN [1]
RX SEQUENCE FROM N.A.
TISSUE-Pancreas.
RA "Regulation of PKA expression in a hibernating mammal."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AY038604; AA072558.1;
CC
CC HSSP: PC1309; 1LNP.
CC
CC InterPro: IPR004825; Ins/IGF/relax.

```

```

Pfam: PF00049; Insulin; 1
PRINTS: P800277; INSULIN;
SMART: SM00028; 106; 1
PROSITE: PS00262; INSULIN; 1
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24 BY SIMILARITY
FT CHAIN 25 54 INSULIN B CHAIN
FT PROPEP 57 87 C PEPTIDE
FT CHAIN 90 110 INSULIN A CHAIN
FT DISULFID 31 96 INTERCHAIN (BY SIMILARITY)
FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY)
FT DISULFID 95 100 BY SIMILARITY
SQ SEQUENCE 110 AA; 12064 MW; 451175866228EE5 CRC64;

Query Match 45.3%; Score 266; DB 1; Length 110;
Best Local Similarity 57.4%; Pred. No. 2,2e-21;
Matches 54; Conservative 1; Mismatches 3; Indels 16; Gaps 2;

QY 50 LGTGP--REVNOHLCGSHLVEALYLVGGERGFFYTPKRSREVEKQGGQVGLARGGPGASL 76
DB 17 LGPDPAQAFVNOHLCGSHLVEALYLVGGERGFFYTPKRSREVEKQGGQVGLARGGPGASL 76
QY 86 -----RIVFOCCSTICSLYLENYCN 107
DB 77 PQPALEMAKQKRGIVFOCCSTICSLYLENYCN 110

RESULT: 9
INS_BALBO . STANDARD; PRT; 51 AA.
AC P01314;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
EN Insulin.
GN INS.
OS Balaenoptera borealis (Sei whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
EX NCBI_TaxID=9768;
RN [1]
RP SEQUENCE.
RA Ishihara Y., Saito T., Ito Y., Fujino M.;
RT "Structure of sperm and sei whale insulins and their breakdown by
RL whale pepsin.";
RL Nature 181:1468-1469(1958).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A01582; INWELS.
DR HSSP: P01317; 1429.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5723 MW; 9007B50F40CA7DDC CRC64;

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 1,9e-21;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Pfam: PF00049; Insulin; 1
PRINTS: P800277; INSULIN;
SMART: SM00028; 106; 1
PROSITE: PS00262; INSULIN; 1
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24 BY SIMILARITY
FT CHAIN 25 54 INSULIN B CHAIN
FT PROPEP 57 87 C PEPTIDE
FT CHAIN 90 110 INSULIN A CHAIN
FT DISULFID 31 96 INTERCHAIN (BY SIMILARITY)
FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY)
FT DISULFID 95 100 BY SIMILARITY
SQ SEQUENCE 110 AA; 12064 MW; 451175866228EE5 CRC64;

Query Match 45.3%; Score 266; DB 1; Length 110;
Best Local Similarity 57.4%; Pred. No. 2,2e-21;
Matches 54; Conservative 1; Mismatches 3; Indels 16; Gaps 2;

QY 50 LGTGP--REVNOHLCGSHLVEALYLVGGERGFFYTPKRSREVEKQGGQVGLARGGPGASL 76
DB 17 LGPDPAQAFVNOHLCGSHLVEALYLVGGERGFFYTPKRSREVEKQGGQVGLARGGPGASL 76
QY 86 -----RIVFOCCSTICSLYLENYCN 107
DB 77 PQPALEMAKQKRGIVFOCCSTICSLYLENYCN 110

RESULT: 9
INS_BALBO . STANDARD; PRT; 51 AA.
AC P01314;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
EN Insulin.
GN INS.
OS Balaenoptera borealis (Sei whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
EX NCBI_TaxID=9768;
RN [1]
RP SEQUENCE.
RA Ishihara Y., Saito T., Ito Y., Fujino M.;
RT "Structure of sperm and sei whale insulins and their breakdown by
RL whale pepsin.";
RL Nature 181:1468-1469(1958).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A01582; INWELS.
DR HSSP: P01317; 1429.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5723 MW; 9007B50F40CA7DDC CRC64;

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 1,9e-21;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 56 FVNQHLCGSHLVEALYLVGGERGFFYTPKTRIGIVEOCCTSCISLYOLENYCN 107
DB 1 FVNQHLCGSHLVEALYLVGGERGFFYTPKTRIGIVEOCCTSCISLYOLENYCN 51

RESULT: 10
INS_CAMER . STANDARD; PRT; 51 AA.
AC P01320;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
EN Insulin.
GN INS.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
EX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RA Danho W.O.;
RT "The isolation and characterization of insulin of camel (Camelus
RL dromedarius).";
RL J. Fac. Med. Baghdad 14:16-28(1972).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A32782; INCM.
DR HSSP: P01317; 21NS.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5693 MW; 901E88BA085A7DDD CRC64;

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 90.4%; Pred. No. 1,9e-21;
Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 56 FVNQHLCGSHLVEALYLVGGERGFFYTPKTRIGIVEOCCTSCISLYOLENYCN 107
DB 1 FVNQHLCGSHLVEALYLVGGERGFFYTPKTRIGIVEOCCTSCISLYOLENYCN 51

RESULT: 11
INS_CAMER . STANDARD; PRT; 51 AA.
AC P01319;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
EN Insulin.
GN INS.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
EX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE.
RA MEDLINE=66160119; PubMed=5949593;
RA Smith L.F.;

```



```

DR PDB: 1ZMJ: 28-JAN-98.
DR PDB: 1ZMJ: 28-JAN-98.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Secreted; 3D structure.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 88 108 INSULIN A CHAIN.
FT DISULFID 91 94 INTERCHAIN.
FT DISULFID 43 107 INTERCHAIN.
FT DISULFID 93 98
FT HELIX 26 44
FT STRAND 48 48
FT STRAND 59 94
FT HELIX 100 106
FT STRAND 107 107
SQ SEQUENCE 108 AA: 11671 MW: 544918429558886 CRC64:

Query Match 44.8%; Score 263; DB 1; Length 106;
Best Local Similarity 60.7%; Pred. NO. 4.6e-21;
Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1;

QY 56 FVNHLCGSHVEALYLVGGRGFYTPKT----- 35
DB 25 FVNHLCGSHVEALYLVGGRGFYTPKARAEAFQAGVHGGTGGLOAALRGTF 34

QY 86 --RGVEQCCISCSLYQLENYCN 107
DB 85 QRGVEQCCISCSLYQLENYCN 108

RESULT 13
INS_RABIT
ID INS_RABIT STANDARD: PRT: 110 AA.
AC F01311;
DT 2-JUL-1996 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 03, Last sequence update)
DT 01-OCT-1996 (Rel. 04, Last annotation update)
DE Insulin precursor.
GN INS.
CS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Pancreas;
RX MEDLINE=94179210; PubMed=8132571;
RA Devaskar S.J., Giddings S.J., Rajakumar P.A., Carnaqui L.R.,
RA Menon R.K., Zahm D.S.;
RT "Insulin gene expression and insulin synthesis in mammalian neuronal
RT cells.";
RJ J. Biol. Chem. 269:8445-8454(1994);
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=66160119; PubMed=5945594;
RA Smith L.P.;
RT "Species variation in the amino acid sequence of insulin.";
RJ Am. J. Med. 40:662-666(1966).
RN [3]
RP SEQUENCE OF 56-110 FROM N.A.
RA Giddings S.J., Carnaqui L.R., Devaskar S.J.;
RP Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
EMBL: U03610; AAAL9C33.1;
EMBL: M61153; AAAL7546.1;
PIR: A53438; INRR.
HSSP: P01308; IIGF.
InterPro: IPR004825; Ins/IGF/relax.
Pfam: PF00049; Insulin; 1.
SMART: SM00078; IIGF; 1.
PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Secreted; 3D structure.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 83 F -> Y (JN REF. 3).
SQ SEQUENCE 110 AA: 11638 MW: 822975385D77F88 CRC64:

Query Match 44.8%; Score 263; DB 1; Length 110;
Best Local Similarity 59.3%; Pred. NO. 4.6e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNHLCGSHVEALYLVGGRGFYTPKT----- 85
DB 25 FVNHLCGSHVEALYLVGGRGFYTPKSREVELOVQAGLGGGAGGLQPSALEL 84

QY 86 ----RGVEQCCISCSLYQLENYCN 107
DB 85 ALOKRGVEQCCISCSLYQLENYCN 110

RESULT 14
INS_FELCA
ID INS_FELCA STANDARD: PRT: 51 AA.
AC P06306;
DT 01-JAN-1986 (Rel. 06, Created)
DT 01-JAN-1986 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
GN INS.
CS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RX MEDLINE=86214376; PubMed=3518635;
RA Hadden G., Gavellin G., Mutt V., Joernvall H.;
RT "Characterization of cat insulin.";
RJ Arch. Biochem. Biophys. 247:20-27(1986).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
PIR: A01588; INCT.
HSSP: P01317; IIGF.
KW InterPro: IPR004825; Ins/IGF/relax.
RP PRINTS: PR00277; INSULINB.

```

```

DR SMART: SMG0078; ILGF: 1.
DR PROSITE: PS00262; INSULIN: 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN H CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT CHAIN 51 52 INTERCHAIN.
FT DISULFID 7 37
FT DISULFID 19 50
FT DISULFID 36 41
SQ SEQUENCE 51 AA: 5745 MW: 906785096AA07D0C CRC64:
Query Match 44.7%; Score 262.5; DP 1; Length 110;
Best Local Similarity 90.3%; Pred. No. 5.9e-21;
Matches 47; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
QY 56 FVNQHLCGSHLVEAIVGVGGFYTPTKRGIVGQVSTSTSLYQLNHCYN 107
DB 1 FVNQHLCGSHLVEAIVGVGGFYTPTKRGIVGQVSTSTSLYQLNHCYN 51
RESULT 15
INS_CANFA
ID INS_CANFA STANDARD: PRI: 110 AA.
AC P01321;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
CS Canis familiaris (Dog).
OC Eukaryote; Metazoa; Chordata; Mammalia; Carnivora; Euteleostomi; Euteleostomi;
OC Mammalia; Euteleostomi; Carnivora; Euteleostomi; Euteleostomi;
CX NCBI_TaxID:9615;
RN [2];
RP SEQUENCE FROM N.A.
RX MEDLINE:83109071; PubMed:6236142;
RA Kwok S.C.M., Chan S.J., Steiner D.F.;
RT "Cloning and nucleotide sequence analysis of the dog insulin gene."
RT Coded amino acid sequence of canine proinsulin predicts an
RT additional C-peptide fragment.
RL J. Biol. Chem. 258:2357-2363(1983).
RN [2];
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE:66160119; PubMed:5945993;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/
CC or send an email to license@sib-sb.ch.)
CC
CC EMBL: V00179; CAA23475.1;
CC PIR: A92413; .PCG.
CC HSSP: P01317; IAPH.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00349; Insulin; 1.
CC SMART: SMG0078; ILGF: 1.
DR PROSITE: PS00262; INSULIN: 1.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24

```

```

FT CHAIN 25 54 INSULIN H CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12190 MW: A574791864A4FB98 CRC64:
Query Match 44.6%; Score 262; DP 1; Length 110;
Best Local Similarity 59.3%; Pred. No. 5.9e-21;
Matches 51; Conservative 0; Mismatches 1; Indels 34; Gaps 1;
QY 56 FVNQHLCGSHLVEAIVGVGGFYTPTKRGIVGQVSTSTSLYQLNHCYN 107
DB 25 FVNQHLCGSHLVEAIVGVGGFYTPTKRGIVGQVSTSTSLYQLNHCYN 85
QY 86 ----RGIVEQCCTSCSLYQLNHCYN 107
DB 85 ALQKRGIVEQCCTSCSLYQLNHCYN 110

```

Search completed: September 16, 2003, 12:38:51
 Job time : 9.32685 secs

Result No.	Query		Length	DB ID	Description
	Score	Match			
1	267	45.5	110	6	Q8XV2
2	266	45.3	110	1	Q91X73
3	251	42.8	110	6	Q9XN6
4	249	42.4	217	6	Q9XN6
5	228	38.8	245	4	Q14644
6	219.5	37.4	106	13	Q91XQ7
7	213	36.3	212	5	Q07368
8	213	36.5	217	5	Q07367
9	201.5	34.3	110	13	Q947A6
10	201	34.2	217	6	Q07369
11	197	33.6	217	4	Q14307
12	195.5	33.3	108	13	Q90CE5
13	195.5	33.3	108	13	Q90ZK4
14	195	33.2	111	13	Q90TB0
15	195	33.2	217	6	Q9XN6
16	193.5	33.0	110	13	Q90ZY1


```

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Insulin.
OS Spermatophytes tridecollineatus (thirteen-lined ground squirrel)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurimorphi; Sciuridae; Spermophiles;
OC Spermophiles;
OX NCBI_TaxID=44179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Tiedra M.M., Buck M.J., Gubanyiogi J., Squire T.L., Andrews M.T.;
RT "Regulation of PK4 expression in a hibernating mammal.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AY038604; AAK72556.1; -.
DR HSSP: P01306; ILNF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12064 MW: 45176806222885 CRC64;

Query Match 45.3%; Score 246; DB 1; Length 110;
Best Local Similarity 57.4%; Pred. No. 8,70-24;
Matches 54; Conservative 1; Mismatches 3; Indels 16; Gaps 2;

QY 50 LGTPG--RVVNHLCGSHVEALYLVCGERGFYTKT----- 85
DB 17 LGPPAQAFVNHLCGSHVEALYLVCGERGFYTKSRREVVEQGGQVFLGSGPGAGL 76

QY 86 -----RGIVEQCCTISGLYOLENYCN 107
DB 77 PPTALEMALKRGIVEQCCTISGLYOLENYCN 110

RESULT 3
Q8WNW6 PRELIMINARY: PRI: 110 AA.
AC Q8WNW6;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Preproinsulin.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Okamoto S., Morimatsu M.;
RT "cat insulin.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AB043535; BAB04116.1; -.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12069 MW: 9546821706738241 CRC64;

Query Match 42.8%; Score 251; DB 6; Length 110;
Best Local Similarity 55.8%; Pred. No. 5,90-22;
Matches 48; Conservative 2; Mismatches 2; Indels 34; Gaps 1;

QY 56 FVNHLCGSHVEALYLVCGERGFYTKT----- 85
DB 25 FVNHLCGSHVEALYLVCGERGFYTKARREEDLQKIALGDAAGGLQPSALEA 84

```

```

QY 86 ----RGIVEQCCTISGLYOLENYCN 107
DB 85 PLOKRGVECCASVSLYOLENYCN 110

RESULT 4
Q8WNE0 PRELIMINARY: PRI: 217 AA.
AC Q8WNE0;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Growth hormone.
OS Gn-N.
OS Ateles geoffroyi (Black-handed spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=9509;
RN [1]
RP SEQUENCE FROM N.A.
RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;
RT "Independent duplication of the growth hormone gene in three
RT Artiodactylid lineages.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF34234; AAL72286.1; -.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24894 MW: 425829944185AAR6 CRC64;

Query Match 42.4%; Score 249; DB 6; Length 217;
Best Local Similarity 97.9%; Pred. No. 20-21;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFNAMLRHRLHQLAFDTYQFEFAYTPKEQKYSFLQNP 49
DB 27 FPTIPLSRLFNAMLRHRLHQLAFDTYQFEFAYTPKEQKYSFLQNP 74

RESULT 5
Q14644 PRELIMINARY: PRI: 245 AA.
AC Q14644;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Placental growth hormone isoform hGH-V3 precursor.
OS HCR-V.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Term placenta;
RA MEDLIN=983737; PubMed=9709563;
RA Bonuszewski C.M., Svensson B.A., Jansson T., Clark R.,
RA Carlsson L.M.S., Carlsson B.;
RL "Cloning of two novel growth hormone transcripts expressed in human
RL placenta.";
RL J. Clin. Endocrinol. Metab. 83:2878-2885(1998).
DR EMBL: AF006061; AAB71829.1; -.
DR HSSP: P0241; IAZ2.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR SIGNAL.
ET SIGNAL.
SQ SEQUENCE 245 AA: 2710; MW: 14007807509;C8 CRC64;

Query Match 38.8%; Score 228; DB 4; Length 245;

```

Best Local Similarity 92.7%; Pred. No. 7,36-137
Matches 44; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

[illegible]

RESULT 6
Q918Q7
ID Q918Q7
PRELIMINARY: PRT: 106 AA.

DT	01-OCT-2000	(TRENALrel. 15, Created)
DT	01-OCT-2000	(TRENALrel. 15, Last sequence update)
DT	01-MAR-2003	(TRENALrel. 23, Last annotation update)

OS Rana pipiens (Northern leopard frog).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana
CX NCBI TaxID-8404.

RP SEQUENCE FROM N.A.
RX MEDLINE-20362507; PubMed-17314274;
RA Irwin D.M., Sivaratnam P.;
RT *proinsulin cDNAs from the leopard frog, *Rana pipiens*, contain both
RT proinsulin processing
RL Comp. Biochem. Physiol. 125B:405-410(2000).
CC -1- SURGEGULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

DR HSSP: P01315; 1SDR.
DR InterPro: IPR004825; Ins/ISF/relax.

DR PRINTS; PROG277; INSULIN3.
DR SMART; SM00078; IGF; 1.

Query Match	Score	Pos	Seq	Sequence	106 AA	12163 N6	3A870EEC70217F92	CR034
37	49	6000	216	5	100	1	1	1

Best Local Similarity 49.4%; Prod. No. 2,805-18;
Matches 41; Conservative 7; Mismatches 4; Indels 3; gaps 1;

55 PUNHLCSSHLVFLVVGEGGFYTPKTR-----AD
 24 PUNGI CCHHVEK VMLTCCPCVNSDSDYKQHLVNI CCELEFSSGALVLPDZG 98

QY 87 --GIVEQCCTSECSLYQLENYCN 107

000 84 AF GIVE QU CHN. CSLSYDIENYCN 100

QC7368 ID Q07368 PRELIMINARY; PRT: 212 AA.

DT	01-NOV-1996 (TREM3.rel. 01, Created)
DT	01-NOV-1996 (TREM3.rel. 01, Last sequence update)

DE Somatotropin 2 precursor (Growth hormone 2) (Fragment).
QS *Macaca mulatta* (Rhesus macaque).

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.

RN	[1]
RP	SEQUENCE FROM N.A.
PC	TESTING OF COAT

RX MEDLINE=94008724; PubMed=8434617;
 RA Golos T.G., Darning M., Fisher J.M., Fowler P.D.;
 PT "Cloning of four growth hormone/chorionic growth hormone-related

RT: complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.";
Endocrinology 133:1744-1752(1993)
RT: Endocrinology 133:1744-1752(1993)

```

DK      EMBL: L16553; AAA18340.1; -.
DK      HSP: PG1241; 1AX1.
DK      InterPro: IPR001400; Somatotropin.
DK      PRAM: PF00103; hormone; 1.
DK      PROSITE: PS00338; SOMATOTROPIN_2; 1.
DK      NON_TER 1
DK      SEQUENCE 212 AA; 24525 MW; 27BC9110625656F5 CRC64;

```

Query Match: 36.3%; Score 213; DB 6; Length 212;
Best Local Similarity 78.7%;
Pred. No. 3.7e-17;
Matches 37; Conservative 9; Mismatches 1; Indels

CY 3 PTIPLSRIFDNAMIRAHRLHQLAFDTYQFEFEAYIPKEQKYSFTQN P 49
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 23 PSVFLSRFLFDHAMIQARLHQLAFDTYQFEFEAYIPKEKKISLMENP 69

RESULT 8
Q07367
ID Q07367
PRELIMINARY: PRT; 217 AA.

DE 01-NOV-1996 (TRENDEL, 01, created)
 DE 01-NOV-1996 (TRENDEL, 01, last sequence update)
 DE 01-MAR-2003 (TRENDEL, 23, last annotation update)
 DE Chelonic soratamacetropin-1.
 DE Kacaca mutata (Rhesus macaque).
 DE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 DE Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea:
 DE Cercopithecoidea: Macaca.
 DE NCBI_TaxID=9544;
 DE NCBI_TaxID=9544;

REF SEQUENCE FROM N.A.
 NC TISSUE=McGrenney Placenta;

FA
Golos T.G., Burnington M., Fisher J.M., Fowler P.D.,
et al. "Cloning of four growth hormone/chorionic somatomatotropin

RT pregnancy in the rhesus monkey placenta.";
Endocrinology 133:1744-1752(1993).

DR HSP; 301241; AXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: pfam103; hormone; 1

[illegible]

Query Match 36.3%; Score 213; DB 5; Length 217;
Best Local Similarity 78.7%; Pred. No. 3.8e-17;
Matches 37; Conservative 9; Mismatches 1; Indels

C7

3 PTIPLSRLLFDNAMIARAHLRHOLAFDTYCEEEAYIPKEQKYSFLQN**P 49**

: :: : : : : : : : : : : : : : : : : :

D

28 PSVELSRLLFDHAMICAIHRLHOLAFDTYCEEEAYIPKEKHSLMENP**P 74**

RESULT 9
99TA8

```

AC  C98TA8;
LT  01-JUN-2001 (TREMBlrel. 17, Created)
DT  01-JUN-2001 (TREMBlrel. 17, Last sequence update)

```

OS panodon buchholzi (Butterflyfish).
DE preproinsulin.
01 MAR 2003 (JREBrefl. 23, last annotation update)

CC Osteoglossiformes; Pantodontidae; Pantodon.

ON	NAME	DATE	BY	REMARKS
RN	[]			
RP	SEQUENCE FROM N.A.			

RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;

```

RT *Molecular cloning of preproinsulin cDNAs from several
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -:- SIMILARITY: METAZOAN: SECRETED (BY SIMILARITY).
CC -:- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
DR EMBL: AF159586; AAK2972.1; -.
DR HSSP: P01408; IHSI.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00649; Insulin_1.
DR SMART: SM0074; IIGF_1.
DR PROSITE: PS00262; INSULIN.
SQ SEQUENCE 110 AA: 1232; MW: 36000.6592872E04 CRC64;

Query Match 34.3%; Score 201.5; DB 1; Length 116;
Best Local Similarity 43.5%; Pred. No. 4e-16;
Matches 37; Conservative 8; Mismatches 5; Indels 35; Gaps 1;

QY 58 NCHLGGSHVPEALVYVSGRGGFFYPKT----- 85
DB 26 SCHLGGSHVLDALVWCGKGGFFCPKTKRVGVNLLAFLEPKSAGENAEEYFKQSL 85

QY 86 ---RGIYRQCCTSCISLYLENYCN 107
DB 86 KVKRGIVEQCCHPCNIFCLQNYCN 110

RESULT 10
Q07369 PRELIMINARY; PRT: 217 AA.
AC Q07369;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chorionic somatomammotropin-3.
OS Macaca mulatta (Rhesus macaque).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: Midpregnancy placenta;
FX MEDLINE=94008724; PubMed=840612;
NA Golos T.G., Durning M., Fisher J.K., Fowler P.D.;
RT "Cloning of four growth hormone/chorionic somatomammotropin-related
RT complementary deoxyribonucleic acids differentially expressed during
RT pregnancy in the rhesus monkey placenta."
RL Endocrinology 133:1744-1752(1993).
DR EMBL: L16554; AAA18841.1;
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1;
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24874; MW: 27686.070615185 CRC64;

Query Match 34.28; Score 201.5; DB 6; Length 217;
Best Local Similarity 74.5%; Pred. No. 1e-15;
Matches 35; Conservative 8; Mismatches 4; Indels 5; Gaps 1;

QY 3 PTFLSRLEFDNIMLAHRLHQAFLAFDTYQEEFEAYIPKQKYSPLQN 49
DB 28 PSVPLSRLEFDNIMQALHQAFLAFDTYQEEFEAYIPKQKYSPLQN 74

RESULT 11
Q14407 PRELIMINARY; PRT: 217 AA.
AC Q14407;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chorionic somatomammotropin CS-2 (Chorionic somatomammotropin hormone-

```

```

DE 21; Homo sapiens (human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=89307277; PubMed=2744760;
FX Chen F.Y., Hsiao Y.C., Smith D.H., Barrera-Saldana H.A., Gollins R.E.,
FX Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
RT evolution."
RL Genomics 4:479-497(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=91102558; PubMed=1980158;
FX Vencsak-Torres C.L., Phillips J.A. III.;
RT "Hot spots for growth hormone gene deletions in homologous regions
RT outside of Alu repeats."
RL Science 250:1745-1748(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
FX Strausberg R.;
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
FX Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: J03071; AAA52553.1; -.
DR EMBL: BC022044; AAH22044.1; -.
DR EMBL: BC035965; AAH35965.1; -.
DR HSSP: P01241; IAX2.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1;
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24994; MW: 39FAACDDB6R2E951 CRC64;

Query Match 43.6%; Score 197; DB 4; Length 217;
Best Local Similarity 80.8%; Pred. No. 3e-15;
Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 TPLSRLEFDNIMLAHRLHQAFLAFDTYQEEFEAYIPKQKYSPLQN 48
DB 29 TVPLSRLEFDNIMLAHRLHQAFLAFDTYQEEFEAYIPKQKYSPLQN 73

RESULT 12
Q07369 PRELIMINARY; PRT: 108 AA.
AC Q07369;
DT 01-MAR-2003 (TrEMBLrel. 16, Created)
DT 01-MAR-2003 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin precursor.
GN INS.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
CC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99425150; PubMed=10495291;
FX Argenton F., Zecchin E., Bortolussi M.;
RT "Early appearance of pancreatic hormone-expressing cells in the
RT zebrafish embryo."
RL Mech. Dev. 87:217-221(1999).
CC -:- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -:- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

```

[illegible]

Db 27 FBIPIPLSRLECHAMLEAHQHQHVANDFYCELEENCIPKKGKXTMKNP 24

Search completed: September 16, 2003, 12:40:00
Job time : 27.8132 secs

Result No.	Query		Length	DP	ID	Description
	Score	Match				
1	797	100.0	150	20	AAV42851	Chimeric protein
2	555.5	69.7	107	20	AAV42850	EGH-minip protein
3	470	59.0	92	20	AAV42856	Human growth hormone
4	470	58.0	134	20	AAV52255	Human anti-angioge
5	470	59.0	191	23	ARG3496	Human growth hormone
6	466	58.5	192	10	AAV52129	Human growth hormone
7	466	58.5	192	20	AAV52264	Human anti-angioge
8	465	58.3	140	10	AAV9104	Human growth hormone
9	465	58.3	261	10	AAV91296	Human growth hormone

PT particularly for the production of human insulin

PS Claim 14; Page 30-31; 46pp; English.

XX This sequence represents a chimeric protein, which contains an

CC N-terminal fragment of human growth hormone (hGH) of the sequence given

CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin

CC precursor comprising insulin A and B chains (AAY42859). The hGH portion

CC of the chimeric protein acts as an intramolecular chaperone (IMC) for

CC the insulin precursor, enabling it to fold correctly. The cleavable

CC peptide linker has a C-terminal Arg residue which enables the hGH portion

CC of the chimeric protein to be removed after folding has taken place. Production

CC of recombinant human insulin via an hGH-proinsulin chimeric protein can

CC provide human insulin with correctly linked cysteine bridges with

CC fewer necessary procedural steps, and hence resulting in a higher yield

CC of human insulin. The IMC sequences not only protect insulin sequences

CC from intracellular degradation by a microorganism host, but also promote

CC the folding of the fused insulin precursor, facilitate the solubility of

CC the fusion protein, and decrease the intermolecular interactions among

CC the fused insulin precursor at commercially useful high concentrations. The

CC procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis

CC and related purification steps can thus be eliminated, along with the use

CC of high concentrations of mercaptan or the use of hydrophobic absorbent

CC resins.

XX

SQ Sequence 150 AA;

Query Match 100.0% Score 797; DB 20; Length 150;

Best Local Similarity 100.0% Pred. No. 10-42;

Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSFSESIP 60

DB 1 MEPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSFSESIP 60

QY 61 TPSNREETOQKSNLELRISLLLIQSWLEPVLGTGPRFVNQHLCCGSHLVEALYLVCGER 120

DB 61 TPSNREETOQKSNLELRISLLLIQSWLEPVLGTGPRFVNQHLCCGSHLVEALYLVCGER 120

QY 121 GFFYTPKTRGIVEQCCTISGLYQLENYCN 150

DB 121 GFFYTPKTRGIVEQCCTISGLYQLENYCN 150

RESULT 2

AAY42860

ID AAY42860 standard; protein; 107 AA.

XX AAY42860;

AC AAY42860;

CT 19-JAN-2000 (first entry)

XX hGH-mini proinsulin chimeric protein.

DE insulin; precursor; growth hormone; chaperone; intramolecular;

KW folding; conformation; chimeric protein; cleavable; recombinant;

KW production; yield.

XX Synthetic.

OS Homo sapiens.

OS WO950302-A1.

PN 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN00052

XX 31-MAR-1998; 98WO-CN00052

PR (TONG-) TONGHUA GANTECH BIO-TECHNOLOGY LTD.

XX Can 2;

PI

XX WPI; 1993-610839/52.

XX New chimeric proteins containing human growth hormone fragment, used

PT particularly for the production of human insulin

XX Claim 13; Page 30; 46pp; English.

XX This sequence represents a chimeric protein, hGH-mini-proinsulin

CC This chimeric protein contains an N-terminal fragment of human growth

CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide

CC linker (AAY42857), and a human insulin precursor comprising insulin

CC A and B chains (AAY42859). The hGH portion of the chimeric protein acts

CC as an intramolecular chaperone (IMC) for the insulin precursor,

CC enabling it to fold correctly. The cleavable peptide linker has a

CC C-terminal Arg residue which enables the hGH portion of the

CC chimeric protein to be removed after folding has taken place. Production

CC of recombinant human insulin via an hGH-proinsulin chimeric protein can

CC provide human insulin with correctly linked cysteine bridges with

CC fewer necessary procedural steps, and hence resulting in a higher yield

CC of human insulin. The IMC sequences not only protect insulin sequences

CC from intracellular degradation by a microorganism host, but also promote

CC the folding of the fused insulin precursor, facilitate the solubility of

CC the fusion protein, and decrease the intermolecular interactions among

CC the fused insulin precursor at commercially useful high concentrations. The

CC procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis

CC and related purification steps can thus be eliminated, along with the use of high

CC concentrations of mercaptan or the use of hydrophobic absorbent resins.

XX

SQ Sequence 107 AA;

Query Match 69.7% Score 555.5; DB 20; Length 107;

Best Local Similarity 71.3% Pred. No. 8e-28;

Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MEPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSFSESIP 60

DB 1 MEPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSFSESIP 60

QY 61 TPSNREETOQKSNLELRISLLLIQSWLEPVLGTGPRFVNQHLCCGSHLVEALYLVCGER 120

DB 61 TPSNREETOQKSNLELRISLLLIQSWLEPVLGTGPRFVNQHLCCGSHLVEALYLVCGER 120

DE 50 -----LGTGPRFVNQHLCCGSHLVEALYLVCGER 77

RESULT 3

AAY42856

ID AAY42856 standard; protein; 92 AA.

XX AAY42856;

AC AAY42856;

CT 19-JAN-2000 (first entry)

XX Human: growth hormone (hGH) N-terminal fragment #2.

DE Growth hormone; chaperone; intramolecular; insulin; precursor;

KW folding; conformation; chimeric protein; cleavable; recombinant;

KW production; yield.

XX Homo sapiens.

OS WO950302-A1.

PN 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN00052.

XX 31-MAR-1998; 98WO-CN00052.

PR (TONG-) TONGHUA GANTECH BIO-TECHNOLOGY LTD.

XX Can 2;

PI

ID AAW92264 Standard; Protein: 192 AA.
 AC AAW92264:
 DT 08-JUN-1999 (first entry)
 DE Human anti-angiogenic peptide hGH Met-1Phc191.
 KW Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
 KW growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
 KW placental vascularisation; pregnancy; treatment; angiogenic disease;
 KW tumour; inhibitor; malignancy; angiofibroma; arteriovenous malformation;
 KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;
 KW wound healing; proliferative retinopathy; macular degeneration; trachoma;
 KW granulation; glaucoma; ocular; uveitis; fracture; Oster-Weber syndrome;
 KW psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
 KW ulcer; leukaemia; reproductive disorder; contraceptive agent;
 KW gene therapy; pre-eclampsia; intrauterine growth retardation;
 KW placental dysfunction.
 OS Homo sapiens.
 XX
 XX W09851323-A1.
 XX
 XX 13-NOV-1998.
 XX
 XX 12-MAY-1998: 98W0-US03691.
 XX
 XX 13-MAY-1997: 97US-0043394.
 XX
 XX (PESC) CNIV CALIFORNIA.
 XX
 XX Martial JA, Struman I, Taylor R, Weiner RJ:
 XX WPI: 1999-045122/0A.
 XX N-PSDB; AAW92706.
 XX
 XX New anti-angiogenic peptides - comprise N-terminal fragments of
 XX human placental lactogen, human growth hormone, growth hormone
 XX variant or human prolactin.
 XX
 XX Example 3: Page 49; 87pp; English.
 XX
 XX This invention describes novel human anti-angiogenic peptides derived
 XX from 10 to 150 consecutive amino acids selected from the N-terminal end
 XX of human placental lactogen (hPL), human growth hormone (hGH), growth
 XX hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
 XX capillary endothelial cell proliferation and organisation (ii) inhibit
 XX angiogenesis in chick chorioallantoic membrane and (iii) binds to at
 XX least one specific receptor which does not bind an intact full length
 XX hGH, hPL, prolactin or hGH-V. The invention also describes a method for
 XX diagnosing a probable abnormality of placental vascularisation during
 XX pregnancy. The peptides can be used for treating an angiogenic disease in
 XX a subject, for inhibiting tumour formation or growth in a patient or for
 XX modulating vascularisation of a patient's placenta. In particular, the
 XX peptides can be used for preventing or treating e.g. malignant tumours,
 XX angiofibroma, arteriovenous malformation, arthritic such as rheumatoid
 XX arthritis, atherosclerotic plaques, corneal graft neovascularisation,
 XX delayed wound healing, proliferative retinopathy such as diabetic
 XX retinopathy, macular degeneration, granulation such as those occurring
 XX in haemophilic joints, inappropriate vascularisation in wound healing
 XX such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular
 XX tumour, uveitis, non-union fractures, Oster-Weber syndrome, psoriasis,
 XX pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,
 XX Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,
 XX leukaemia, and reproductive disorders such as follicular and luteal cysts
 XX and choriocarcinoma. They can also be used as contraceptive agents, DNA
 XX encoding the peptides can be used in gene therapy. The measurement of
 XX abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL
 XX can be used in assays for impairment of vascular development associated
 XX with pre-eclampsia, intrauterine growth retardation, and placental
 XX dysfunction.

SV Sequence 192 AA:
 Query Match 58.5%; Score 466; DH 20; Length 192;
 Best Local Similarity 70.5%; Pred. No. 4.6e-22;
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 1 MFTPIPLSLRFDNAMLRAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 DB 1 MFTPIPLSLRFDNAMLRAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 QY 61 TPSNRETQCKSNLELLRISLLLIQSWLEPVLGICGPRFVNOHLCGS-----HGV 110
 DB 61 TPSNRETQCKSNLELLRISLLLIQSWLEPVLGICGPRFVNOHLCGS-----HGV 110
 QY 111 EALYLVCG--ERGFYTPKTRGIVEQ 134
 DB 120 EGIOTLMRHLFDG---SPRTGQIFKQ 142
 RESULT 8
 AAW91041
 ID AAW91041 standard; protein: 140 AA.
 XX
 XX AAW91041:
 XX 14-DEC-1989 (first entry).
 XX Human growth hormone segment.
 XX
 XX Human growth hormone; fusion protein; thrombin;
 XX geriatric dementia; nervous disorders; human nerve factor.
 XX
 XX Homo sapiens (human).
 XX
 XX EP329175-A.
 XX 23-APR-1989.
 XX 17-FEB-1989: 89EP-0102795.
 XX 19-FEB-1982: 88JP-0015042.
 XX (TOYT) TOSOH CORP.
 XX Ohtsuka E.
 XX WPI: 1989-243092/34.
 XX New human nerve growth factor gene encoding fusion protein
 XX - having cleavage site for thrombin, useful for treating geriatric
 XX dementia, etc.
 XX Disclosure: page 21; 38pp; English.
 XX Human growth hormone segment, used at the N-terminal of a fusion
 XX protein, which contains a thrombin recognition site, and human beta nerve
 XX growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to
 XX control geriatric dementia and other nervous disorders and can be
 XX released from the fusion protein by incubation with thrombin (see
 XX AAW90577-9, AAW91034, AAW91299).
 SV Sequence 140 AA:
 Query Match 58.3%; Score 465; DB 10; Length 140;
 Best Local Similarity 98.9%; Pred. No. 4.2e-22;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MFTPIPLSLRFDNAMLRAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 DB 1 MFTPIPLSLRFDNAMLRAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 QY 61 TPSNRETQCKSNLELLRISLLLIQSWLEPVLGICGPRFVNOHLCGS-----HGV 110

PA (TOSC) TOSCH CORP.
 XX
 DR WPI: 1990-062207/09.
 DR N-PSDB: AAO02028.
 XX
 PT Prepn. of human B-cell differentiation factor - from specified DNA
 PT sequence segment. By recombinant DNA technique, gives protein of
 PI specified amino acid sequence.
 XX
 PS Claim 31; Page 9; 17pp; Japanese.
 XX
 CC The protein is produced by fusing DNA encoding BDF (1-7) with DNA
 CC encoding BDF-2 (11-5) and ligating the product into an expression vector
 CC See also AAR05311 and AAR05313.
 XX
 SQ Sequence 310 AA;
 Query Match 58.3%; Score 465; DB 1; Length 400;
 Best Local Similarity 98.9%; Pred. No. 7.7e-22;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MFPTPLSRFDNAMLRAHRLRQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 Db 1 MFPTPLSRFDNAMLRAHRLRQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 QY 61 TPSNREETQCKSNLELRISLLTQSWLEPVQ 92
 Db 61 TPSNREETQCKSNLELRISLLTQSWLEPVQ 92
 RESULT 12
 ABG31862
 ID ABG31862 standard; Protein: 191 AA.
 AC ABG31862;
 DI 05-NOV-2002 (first entry)
 XX Mature human growth hormone (hGH), mutant #4.
 DE Human growth hormone; hGH; Turner's syndrome; achondroplasia;
 KW growth hormone deficiency in adults; GHDA; chronic renal insufficiency;
 KW renal failure in children; acquired immune deficiency syndrome; AIDS;
 KW AIDS wasting; cachexia; mutant; mutagen.
 XX Homo sapiens.
 QS Synthetic.
 XX
 PH Key Location/Qualifiers
 PT Misc-difference 134 /note= "W10 type Arg substituted by Lys"
 XX
 XX W0200255532-A2.
 XX 18-JUL-2002.
 XX 10-JAN-2002; 2002WO-DK00017.
 XX 11-JAN-2001; 2001DK-090042.
 XX 11-JAN-2001; 2001US-26141P.
 XX (MAXY-) MAXYGEN APS.
 XX (MAXY-) MAXYGEN HOLDINGS LTD.
 XX Andersen KV, Drustup J, Christiansen J;
 WPI: 2002-608345/65.
 XX New conjugates exhibiting growth hormone activity, useful for treating
 PT a disease or for manufacturing a medicament for treating a disease.
 PT e.g. Turner's syndrome, growth hormone deficiency, achondroplasia, AIDS
 PT wasting or cachexia -
 XX

ES Claim 10; Page 7; 74pp; English.
 XX
 CC The invention relates to new conjugates, which exhibit growth hormone
 CC (GH) activity and comprise at least one non-polypeptide group covalently
 CC attached to a GH polypeptide. The amino acid sequence of the conjugates
 CC differs from that of wild type human GH in at least one introduced and at
 CC least one removed amino acid residue comprising an attachment group for
 CC the first non-polypeptide group. The conjugate or pharmaceutical
 CC composition is useful for treating a disease or for manufacturing a
 CC medicament for treating a disease, e.g. Turner's syndrome, GH deficiency
 CC in adults (i.e. GHDA), achondroplasia, chronic renal insufficiency or
 CC failure (including renal failure in children), acquired immune deficiency
 CC syndrome (AIDS) wasting, cachexia in AIDS patients, or cachexia
 CC associated with other diseases. The conjugates are useful for treating a
 CC variety of disorders caused by growth hormone inadequacy. The present
 CC sequence represents the amino acid sequence of a mutant human growth
 CC hormone.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the wild type human growth hormone sequence given in SEQ ID
 CC No.2 (see ABG31857).
 XX
 SQ Sequence 191 AA;
 Query Match 58.2%; Score 464; DB 23; Length 191;
 Best Local Similarity 71.0%; Pred. No. 6.1e-22;
 Matches 103; Conservative 6; Mismatches 20; Indels 16; Gaps 4;
 QY 2 FPIPLSRFLFDNAMLRAHRLRQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 61
 Db 2 FPIPLSRFLFDNAMLRAHRLRQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 QY 62 PSNREETQCKSNLELRISLLTQSWLEPVQ 134
 Db 62 PSNREETQCKSNLELRISLLTQSWLEPVQ 134
 QY 112 AUYVQGG--ERGFEYTPKTRGIVEQ 134
 Db 120 GUTLMGRLEDG---SPKTOGEQ 141
 RFSMT 13
 ABG-4860
 IF ABG54860 standard; Protein: 191 AA.
 XX ABG54860;
 AC ABG54860;
 DI 03-DEC-2002 (first entry)
 XX Human growth hormone mutant hPL (109-112).
 XX Growth hormone; placental lactogen; prolactin; active domain; hGH;
 KW structure-function relationship; segment-substituted polypeptide;
 KW mutant; mutagen.
 XX Homo sapiens.
 QS Synthetic.
 XX US6428954-B1.
 XX 06-AUG-2002.
 XX 06-JUN-1995; 95US-0483039.
 XX 26-OCT-1989; 89US-0428066.
 XX 27-APR-1992; 92US-0875204.
 XX 13-OCT-1992; 92US-0960227.
 XX 02-FEB-1994; 94US-0190723.
 XX 28-OCT-1988; 88US-0264611.
 XX (GETH) GENENTECH INC.
 XX Wells JA, Cunningham BC;
 XX

DR WPI: 2002-696875/75.
 XX Identifying active domains within cloned polypeptides of known amino
 PT acid sequence by substituting analog segments into the parent
 PT polypeptide is useful to determine the relationship between structure
 PT and function.
 XX Example 1: Page -: 86pp; English.
 PS
 XX The invention relates to identifying an unknown active domain in a region
 CC of known amino acid sequence in a parent polypeptide (e.g. human growth
 CC hormone (hGH) which has been cloned and has a pre-identified biological
 CC activity, where the active domain interacts with a target when the parent
 CC polypeptide is in its native-folded form and the interaction is
 CC responsible for the biological activity comprising: (a) comparing the
 CC amino acid sequence of polypeptide structure in the region of known amino
 CC acid sequence of hGH with the amino acid sequence of polypeptide
 CC structure in a region of known amino acid sequence of an analog
 CC polypeptide (e.g. prolactin, placental lactogen or porcine growth
 CC hormone) which has at least 15% homology with hGH alpha-carbon
 CC coordinates within about 2-3.5 angstroms of hGH alpha-carbon
 CC for about 60% of the analogue sequence, where any interaction of the
 CC analogue with the target is different from target interaction with hGH;
 CC (b) substituting DNA encoding an analogous polypeptide segment from the
 CC analogue into DNA encoding the full length hGH, and expressing a
 CC segment-substituted polypeptide; (c) contacting the segment-substituted
 CC polypeptide with the target to determine interaction; (d) repeating steps
 CC (b) and (c) with a second analogous polypeptide segment; and
 CC (e) comparing the difference between activity of the first and second
 CC segment-substituted polypeptides as an indication of the location of
 CC the unknown active domain in hGH. The method is useful for determining
 CC the relationship between structure and function of known polypeptide
 CC sequences. The present sequence is that of human growth hormone
 CC mutant substituted with residues from an hGH analogue (prolactin).
 CC Note: The present sequence is not shown in the specification but was
 CC created by the indexer using the mature hGH sequence and information
 CC contained in the specification.
 XX Sequence 191 AA:

Query Match 58.1%; Score 463; DB 23; Length 191;
 Best Local Similarity 71.0%; Pred. No. 7.7e-22;
 Matches 103; Conservative 3; Mismatches 16; Indels 16; Gaps 5;
 QY 2 PFTIPLSRLFDNMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSSEIPT 61
 DB 1 PFTIPLSRLFDNMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSSEIPT 60
 QY 62 PSNREETQCKSNLELLRLSLILIQSWLEPVOIGTPREVNQHLGCS-----HLV-----E 111
 DB 61 PSNREETQCKSNLELLRLSLILIQSWLEPVOF-LRSVPANSLVAGASDNVYDLKOLEE 119
 QY 112 ALYLVCG--ERGFYTPKIRGIVEQ 134
 DB 120 GLOTLMKRLKEDG---SVRTSQIFKQ 141

RESULT 14
 ABC94977
 ID ABC94977 standard: Protein; 191 AA.
 XX
 AC ABC94977;
 XX

CT 03-DEC-2002 (first entry)
 XX
 DE Human growth hormone mutant with 3A.
 XX
 KW Growth hormone; placental lactogen; prolactin; active domain; hGH;
 KW structure-function relationship; segment-substituted polypeptides;
 KW mutant; mutain.
 XX
 OS Homo sapiens.

Synthetic.
 XX US6428954-B1.
 XX 06-AUG-2002.
 XX 06-JUN-1995; 9505 0481037.
 XX 26-OCT-1989; 8905-0426066.
 XX 27-APR-1992; 923S-0852234.
 XX 13-OCT-1992; 923S-0960227.
 XX 02-FEB-1994; 940S-0190723.
 XX 28-OCT-1993; 880S-0264611.
 XX (GETH) GENENTECH INC.
 XX Wells JA. Cunningham BC;
 XX WPI: 2002-696875/75.
 PT Identifying active domains within cloned polypeptides of known amino
 PT acid sequence by substituting analog segments into the parent
 PT polypeptide is useful to determine the relationship between structure
 PT and function.
 XX Example 16: Page -: 86pp; English.
 PS The invention relates to identifying an unknown active domain in a region
 CC of known amino acid sequence in a parent polypeptide (e.g. human growth
 CC hormone (hGH) which has been cloned and has a pre-identified biological
 CC activity, where the active domain interacts with a target when the parent
 CC polypeptide is in its native-folded form and the interaction is
 CC responsible for the biological activity comprising: (a) comparing the
 CC amino acid sequence of polypeptide structure in the region of known amino
 CC acid sequence of hGH with the amino acid sequence of polypeptide
 CC structure in a region of known amino acid sequence of an analog
 CC polypeptide (e.g. prolactin, placental lactogen or porcine growth
 CC hormone) which has at least 15% homology with hGH alpha-carbon
 CC coordinates within about 2-3.5 angstroms of hGH alpha-carbon
 CC for about 60% of the analogue sequence, where any interaction of the
 CC analogue with the target is different from target interaction with hGH;
 CC (b) substituting DNA encoding an analogous polypeptide segment from the
 CC analogue into DNA encoding the full length hGH, and expressing a
 CC segment-substituted polypeptide; (c) contacting the segment-substituted
 CC polypeptide with the target to determine interaction; (d) repeating steps
 CC (b) and (c) with a second analogous polypeptide segment; and
 CC (e) comparing the difference between activity of the first and second
 CC segment-substituted polypeptides as an indication of the location of
 CC the unknown active domain in hGH. The method is useful for determining
 CC the relationship between structure and function of known polypeptide
 CC sequences. The present sequence is that of human growth hormone
 CC mutant substituted at functionally important residues and used in
 CC the method of the invention.
 CC Note: The present sequence is not shown in the specification but was
 CC created by the indexer using the mature hGH sequence and information
 CC contained in the specification.
 XX Sequence 191 AA:

Query Match 58.1%; Score 463; DB 23; Length 191;
 Best Local Similarity 70.3%; Pred. No. 7.7e 22;
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 2 PFTIPLSRLFDNMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSSEIPT 61
 DB 1 PFTIPLSRLFDNMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSSEIPT 60
 QY 62 PSNREETQCKSNLELLRLSLILIQSWLEPVOIGTPREVNQHLGCS-----HLVE 111
 DB 61 PSNREETQCKSNLELLRLSLILIQSWLEPVOF-LRSVPANSLVAGASDNVYDLKOLEE 119
 QY 112 ALYLVCG--ERGFYTPKIRGIVEQ 134
 DB 120 GLOTLMKRLKEDG---SVRTSQIFKQ 141

```

DB      120 GIQTLMGLEDG--SPNIGQIFKQ 141

RESULT 15
AAR05313
ID      AAR05313 standard: protein; 144 AA.
XX
AC      AAR05313;
XX
DI      19-JUL-1990 (first entry)
XX
DE      Segment of B-cell stimulatory factor-2 (IL-5).
XX
KW      B-cell stimulatory factor-2; interleukin-5.
XX
OS      Homo sapiens.
XX
PN      JP02013375-A.
XX
PD      17-JAN-1990.
XX
PF      01-JUL-1988; 8BJP-0162556
XX
PR      01-JUL-1988; 8BJP-0162556
XX
PA      (TOX) ) TOSOH CORP.
XX
WP1: 1990-062207/39.
DR      N-PSDH: AACC028.
XX
XX      Prep. of human B-cell differentiation factor - from specified DNA
PT      sequence segment, by recombinant DNA technique, gives protein at
PT      specified amino acid sequence.
XX
PS      Disclosure; Page 9; 17pp; Japanese.
XX
XX      The sequence encoding this protein can be fused with DNA encoding B-cell
CC      differentiation factor (IL-6) and ligated into an expression vector for
CC      prodn. of a fusion protein.
CC      See also AAR0311.
XX
SQ      Sequence 14; AA.

Query Match 58.6%; Score 452; DB 11; Length 144;
Best Local Similarity 87.8%; Pred. No. 6,60-22;
Matches 93; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 MEPTFLSKLPNAMLRAHRLHQLAFETYQFEFEAYIFKQKYSFLQNTSLSSSESIF 60
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      1 MEPTFLSKLPNAMLRAHRLHQLAFETYQFEFEAYIFKQKYSFLQNTSLSSSESIF 60
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY      61 TPSNREETQOKSNLELRISLLIQSWLEPVQ 92
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      61 TPSNREETQOKSNLELRISLLIQSWLEPVQ 92

```

Search completed: September 16, 2003, 12:38:27
 Job time : 45.358 secs

GenCore version 5.1.1f
Copyright (c) 1993 - 2003 CompuLink Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:37:10 : Search time 23.93 seconds
(without alignments)
255.217 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPTIPISRLFDNAIPARR.....IVEQNTSISLVQENYCN 150

Scoring table: BL/USM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 128717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA: *
- 1: /can2_6/pdata//iaa/VA/GMB.pept *
- 2: /can2_6/pdata//iaa/5R/GMB.pept *
- 3: /can2_6/pdata//iaa/6A/GMB.pept *
- 4: /can2_6/pdata//iaa/6B/GMB.pept *
- 5: /can2_6/pdata//iaa/PCMS/GMB.pept *
- 6: /can2_6/pdata//iaa/backfiles1.pept *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being plotted and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466	58.5	192	1	US-08-093-383-1
2	461	57.8	191	4	US-09-284-878-5
3	461	57.8	194	2	US-08-381-621-4
4	461	57.8	194	3	US-08-459-976-4
5	461	57.8	217	3	US-08-589-028-10
6	461	57.8	217	3	US-08-784-582-19
7	461	57.8	217	3	US-08-784-582-12
8	461	57.8	217	3	US-08-789-626-11
9	461	57.8	217	4	US-09-294-878-1
10	461	57.8	241	4	US-09-424-6205-25
11	461	57.8	245	4	US-09-230-030-66
12	461	57.8	274	3	US-08-784-582-73
13	461	57.8	360	3	US-08-784-582-73
14	455	57.1	191	4	US-09-465-461-1
15	455	57.1	217	1	US-08-187-756C-4
16	455	57.1	217	1	US-08-469-486-51
17	455	57.1	217	2	US-08-459-653-52
18	455	57.1	217	2	US-08-710-324A-4
19	455	57.1	217	4	US-09-411-657-4
20	454	57.0	460	4	US-09-420-810-37
21	454	57.0	461	4	US-09-420-819-36
22	448	56.2	191	3	US-08-800-215C-18
23	446	56.0	191	3	US-08-800-215C-16
24	446	56.0	191	3	US-08-800-215C-20
25	365.5	45.9	176	3	US-08-791-728-1
26	365.5	45.9	176	3	US-08-990-774-1
27	359.5	45.1	176	3	US-08-791-728-2

24	359.5	45.1	176	4	US-08-990-774-2	Sequence 2, Appli
25	343.5	43.0	166	3	5424199-3	Patent No. 5424199
26	334.5	42.0	198	1	US-08-187-756C-5	Sequence 5, Appli
27	334.5	42.0	198	2	US-08-710-324A-5	Sequence 5, Appli
28	334.5	42.0	198	4	US-09-411-657-5	Sequence 5, Appli
29	316.5	39.7	36	1	US-08-160-376A-5	Sequence 5, Appli
30	316.5	39.7	36	1	US-08-383-487-8	Sequence 8, Appli
31	314	39.4	137	3	US-08-400-256-39	Sequence 39, Appli
32	309.5	38.8	146	3	US-08-975-385-48	Sequence 48, Appli
33	309.5	38.8	146	3	US-08-975-365-48	Sequence 48, Appli
34	309	38.8	145	1	US-08-400-256-45	Sequence 45, Appli
35	309	38.8	145	3	US-08-975-385-45	Sequence 45, Appli
36	306.5	38.5	191	1	US-08-468-824-8	Sequence 8, Appli
37	306.5	38.5	191	2	US-09-105-651-1	Sequence 1, Appli
38	304.5	38.2	191	1	US-07-963-331D-4	Sequence 4, Appli
39	304	38.1	63	1	US-08-160-376A-6	Sequence 6, Appli
40	301.5	37.8	190	1	US-08-388-267C-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-093-383-1
Sequence 1, Application US/38093483
Patent No. 5489529
GENERAL INFORMATION:
APPLICANT: DeHoer, Herman A.
APPLICANT: Heyneker, Herbert L.
APPLICANT: Secburg, Peter H.
TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,383
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/619827
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/198824
FILING DATE: 05-APR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/632361
FILING DATE: 19-JUL-1984
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/303687
FILING DATE: 18-SEP-1981
ATTORNEY/AGENT INFORMATION:
NAME: Johnston, Sean A.
REGISTRATION NUMBER: P35,910
REFERENCE/DOCKET NUMBER: 4604
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3562
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid

TOPOLOGY: linear

US-08-093-383-1

Query Match 58.5% Score 465 DB 1 Length 192

Best Local Similarity 70.3% Pred. No. 3.8e-42

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 60

DB 1 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 60

QY 61 TPSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

DB 61 TPSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

QY 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134

DB 120 GIQTLMGRLDQ---SPRTGQIFKQ 142

RESULT 2

US-09-284-878-5

Sequence 5, Application US/03284878

Patent No. 6342375

GENERAL INFORMATION:

APPLICANT: Olazaran, Martha Guetero

APPLICANT: Saldaña, Rosa Bariera

APPLICANT: Salvado, Jose Maria Viader

TITLE OF INVENTION: Genetically Modified Methylothrophic F. pastorianus Yeast for the

TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone

FILE REFERENCE: 1829.0016000

CURRENT APPLICATION NUMBER: US/09/284.878

CURRENT FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: PCI/WX57/00033

PRIOR FILING DATE: 1997-10-24

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 5

LENGTH: 192

TYPE: PRT

ORGANISM: Homo sapiens

US-09-284-878-5

Query Match

Best Local Similarity 57.8% Score 461 DB 4 Length 191

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 61

DB 3 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 60

QY 62 PSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

DB 61 PSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

QY 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134

DB 120 GIQTLMGRLDQ---SPRTGQIFKQ 142

RESULT 3

US-08-383-621-4

Sequence 4, Application US/08384621

Patent No. 5951972

GENERAL INFORMATION:

APPLICANT: Daley, Michael J.

APPLICANT: Buckwalter, Brian L.

APPLICANT: Cady, Susan M.

APPLICANT: Shieh, Hong-Ming

APPLICANT: Bohlen, Peter

APPLICANT: Seddon, Andrew P.

TITLE OF INVENTION: Stabilization of Somatotropins And Other

TITLE OF INVENTION: Proteins By Modification of Cysteine Residues

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dr. Estelle J. Tsevdos

STREET: 193 West Main Street, P.O. Box 60

CITY: Stamford

STATE: Connecticut

COUNTRY: U.S.A.

ZIP: 06304-0060

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/383.621

FILING DATE: 26-FEB-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/746,142

FILING DATE: 25-SEP-1991

ATTORNEY/AGENT INFORMATION:

NAME: Tsevdos, Estelle J.

REGISTRATION NUMBER: 31,145

REFERENCE/DOCKET NUMBER: 31,278-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 203-321-2756

TELEFAX: 203-321-2971

TELEX: 203-710-474-4959

INFORMATION FOR SEQ ID NO: 4

SEQUENCE CHARACTERISTICS:

LENGTH: 194 amino acids

TYPE: amino acid

TOPOLOGY: linear

MCLEULE TYPE: protein

US-08-383-621-4

Query Match 57.8% Score 461 DB 2 Length 194

Best Local Similarity 70.3% Pred. No. 3.8e-42

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 61

DB 4 MPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESPT 63

QY 62 PSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

DB 64 PSNRETOQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111

QY 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134

DB 123 GIQTLMGRLDQ---SPRTGQIFKQ 144

RESULT 4

US-08-459-906-4

Sequence 4, Application US/08459906

Patent No. 6010959

GENERAL INFORMATION:

APPLICANT: Daley, Michael J.

APPLICANT: Buckwalter, Brian L.

APPLICANT: Cady, Susan M.

APPLICANT: Shieh, Hong-Ming

APPLICANT: Bohlen, Peter

APPLICANT: Seddon, Andrew P.

TITLE OF INVENTION: Stabilization of Somatotropins and Other

TITLE OF INVENTION: Proteins by Modification of Cysteine Residues

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: American Cyanamid Company

STREET: One Cyanamid Plaza

CITY: Wayne

STATE: New Jersey

COUNTRY: U.S.A.


```

1 ATTORNEY/AGENT INFORMATION:
2 NAME: Highlander, Steven L.
3 REGISTRATION NUMBER: 47,542
4 REFERENCE/DOCKET NUMBER: US-0426\HYL
5 TELECOMMUNICATION INFORMATION:
6 TELEPHONE: (512) 418-3000
7 TELEFAX: (512) 474-7577
8 INFORMATION FOR SEQ ID NO: 10:
9 SEQUENCE CHARACTERISTICS:
10 LENGTH: 217 amino acids
11 TYPE: amino acid
12 STRANDEDNESS:
13 TOPOLOGY: linear
14 US-08-585-028-10
15
16 Query Match 57.8%; Score 461; DB 3; Length: 217;
17 Best Local Similarity 70.3%; Pred. No. 4, 4e-42;
18 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4
19
20 QY 2 PFTPLSLFVNMLRAHRLHQLAFDYQFEEAYIPKEGKYSELPQNPTSLSFSESIPT 61
21 DE 27 PFTPLSLFVNMLRAHRLHQLAFDYQFEEAYIPKEGKYSELPQNPTSLSFSESIPT 86
22 QY 62 PSNREETQOKSNLELRISLLIGSNLEPVGQF-LRSVEANSLVYGSDSNVYDLKDL 111
23 DE 87 PSNREETQOKSNLELRISLLIGSNLEPVGQF-LRSVEANSLVYGSDSNVYDLKDL 145
24 QY 112 MLYLCVG--ENGFFYTPKTRGIVEQ 134
25 DE 146 GIQLMGRLDG---SPRTGQIFKQ 167
26
27 RESULT 5
28 US-08-784-542-10
29 Sequence 10, Application US/08784582
30 Patent No. 6110707
31 GENERAL INFORMATION:
32 APPLICANT: Newgard, Christopher B.
33 APPLICANT: Halban, Philippe A.
34 APPLICANT: No. 6110707mington, Karl D.
35 APPLICANT: Clark, Samuel A.
36 APPLICANT: Thigpen, Anice E.
37 APPLICANT: Quade, Christian
38 APPLICANT: Kruse, Fred
39 APPLICANT: McGarry, Dennis
40 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
41 SEQUENCE OF SEQUENCES: 79
42 CORRESPONDENCE ADDRESS:
43 ADDRESSEE: Arnold, White & Durkee
44 STREET: P.O. Box 4433
45 CITY: Houston
46 STATE: Texas
47 COUNTRY: USA
48 ZIP: 77210
49
50 COMPUTER READABLE FORM:
51 MEDIUM TYPE: Floppy disk
52 OPERATING SYSTEM: PC DOS/MS-DOS
53 SOFTWARE: Patent Release #1.0, Version #1.30
54 CURRENT APPLICATION DATA:
55 APPLICATION NUMBER: US/08/784,582
56 FILING DATE: Concurrently Herewith
57 CLASSIFICATION: 435
58 PRIOR APPLICATION DATA:
59 APPLICATION NUMBER: US 60/028,427
60 FILING DATE: 15-OCT-1996
61 PRIOR APPLICATION DATA:
62 APPLICATION NUMBER: US 08/589,028
63 FILING DATE: 19-JAN-1996
64 ATTORNEY/AGENT INFORMATION:
65 NAME: Highlander, Steven L.
66 REGISTRATION NUMBER: 37,642

```

```

: REFERENCE/DOCKET NUMBER: UTSD:514
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 512/418-3000
: TELEFAX: 512/474-7577
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 217 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
US-08-784-582-10

```

```

Query Match          57.8%; Score 461; DB 3; Length 217;
Best Local Similarity 70.3%; Pred. No. 4.4e-42;
Matches 102; Conservative 7; Mismatches 20; Gaps 4;

```

```

QY 2 FTPTLSRLFDNAMLRAHRLHQLADFTYOEFEAYIPKEKYSLQNPQTSISFSESPT 61
DB 27 FTPTLSRLFDNAMLRAHRLHQLADFTYOEFEAYIPKEKYSLQNPQTSISFSESPT 86
QY 62 PSNRETCOKSNLELRISLLIQSKLEPVGQGTGPRFVNHCLGS-----HLVE 111
DB 87 PSNRETCOKSNLELRISLLIQSKLEPVGQGTGPRFVNHCLGS-----HLVE 145
QY 112 ALYLVG--EKGFTYTPKRGIVEQ 134
DB 146 GIOTLMGRLEHG---SPRTGQIFKQ 167

```

```

RESULT 7
US-08-785-271-10
: Sequence 10; Application US/08785271
: Patent No. 6194176
: GENERAL INFORMATION:
: APPLICANT: Newgard, Christopher B.
: APPLICANT: Halbard, Philippe A.
: APPLICANT: No. 6194176;ington, Kar: B.
: APPLICANT: Clark, Samuel A.
: APPLICANT: Thigpen, Anice E.
: APPLICANT: Quaaide, Christian
: APPLICANT: Kruse, Fred
: TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESS: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/785,271
: FILING DATE: Concurrently Herewith
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/569,028
: FILING DATE: 19-2AN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Highlander, Steven L.
: REGISTRATION NUMBER: 37,642
: REFERENCE/DOCKET NUMBER: UTSD:513
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 512/418-3000
: TELEFAX: 512/474-7577
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 217 amino acids

```

```

: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
US-08-785-271-10

```

```

Query Match          57.8%; Score 461; DB 3; Length 217;
Best Local Similarity 70.3%; Pred. No. 4.4e-42;
Matches 102; Conservative 7; Mismatches 20; Gaps 4;

```

```

QY 2 FTPTLSRLFDNAMLRAHRLHQLADFTYOEFEAYIPKEKYSLQNPQTSISFSESPT 61
DB 27 FTPTLSRLFDNAMLRAHRLHQLADFTYOEFEAYIPKEKYSLQNPQTSISFSESPT 86
QY 62 PSNRETCOKSNLELRISLLIQSKLEPVGQGTGPRFVNHCLGS-----HLVE 111
DB 87 PSNRETCOKSNLELRISLLIQSKLEPVGQGTGPRFVNHCLGS-----HLVE 145
QY 112 ALYLVG--EKGFTYTPKRGIVEQ 134
DB 146 GIOTLMGRLEHG---SPRTGQIFKQ 167

```

```

RESULT 8
US-08-759-626-11
: Sequence 11; Application US/08759628
: Patent No. 6223446
: GENERAL INFORMATION:
: APPLICANT: Altmann, Scott W.
: APPLICANT: Rock, Ferrand L.
: APPLICANT: Bazan, J. Fernando
: APPLICANT: Kastelein, Robert A.
: TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: UNAX Research Institute
: STREET: 901 California Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
: ZIP: 94304-1104

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/759,628
: FILING DATE: 05-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/008,574
: FILING DATE: 06-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Ching, Edwin P.
: REGISTRATION NUMBER: 34,090
: REFERENCE/DOCKET NUMBER: DX05520
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-852-9196
: TELEFAX: 415-496-1200
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 217 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Peptide
: LOCATION: 32..53
: FEATURE:
: NAME/KEY: Peptide
: LOCATION: 94..115
: FEATURE:

```

NAME/KEY: Peptide
 LOCATION: 133-153
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 192-210
 OTHER INFORMATION: /note: "the peptides above are
 OTHER INFORMATION: depicted in Figure 1"

US-08-759-628-11

Query Match 57.8%; Score 461; DB 3; Length 217;
 Best Local Similarity 70.3%; Pred. No. 4.4e-42;
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 2 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61
 DB 27 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 60
 QY 62 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 111
 DB 87 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 145
 QY 112 ALYVCG--ERGFYTPKRGIVEQ 134
 DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 9

US-09-284-878-1

Sequence 1; Application US/C92/4978
 Patent No. 6342375
 GENERAL INFORMATION:
 APPLICANT: Olazaran, Martha Guetiero
 APPLICANT: Saldaña, Hugo Barrera
 APPLICANT: Salvador, Jose Maria Viader
 TITLE OF INVENTION: Genetically Modified Melanotrophic P. pastoris Host for the
 PRODUCTION AND SECRETION OF THE Human Growth Hormone
 FILE REFERENCE: 1829.3010000
 CURRENT FILING DATE: 1999-07-21
 PRIOR APPLICATION NUMBER: US/09/284.878
 CURRENT FILING DATE: 1999-07-21
 PRIOR FILING DATE: 1997-10-24
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO: 1
 LENGTH: 217
 TYPE: PRT
 ORGANISM: Homo sapiens

US-09-284-878-1

Query Match 57.8%; Score 461; DB 4; Length 217;
 Best Local Similarity 70.3%; Pred. No. 4.4e-42;
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 2 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61
 DB 27 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 60
 QY 62 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 111
 DB 87 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 145
 QY 112 ALYVCG--ERGFYTPKRGIVEQ 134
 DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 10

US-09-424-620B-25
 Sequence 25; Application US/C94/42620B
 Patent No. 6391585
 GENERAL INFORMATION:
 APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
 JANG, KI-RYONG

NAME, Jae-Woong
 HAE, Cheon-Soon
 YANG, Doo-Suk
 LEE, Jee-Wan
 SEONG, Baik-Lin
 TITLE OF INVENTION: Process for preparing recombinant proteins using high-
 efficient expression vector from Saccharomyces cerevisi.
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BACHMAN & LAPORTE, P.C.
 STREET: Suite 1201, 900 Chapel Street
 CITY: New Haven
 STATE: Connecticut
 COUNTRY: U.S.A.
 ZIP: 06510-2802
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 COMPUTER: IBM
 OPERATING SYSTEM: WINDOWS 95/98
 SOFTWARE: MS WORD
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/424,620B
 FILING DATE: 24-Nov. 6391585-1999
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 241 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: PROTEIN
 SEQUENCE DESCRIPTION: SFO ID NO: 25:
 US-09-424-620B-25

Query Match 57.8%; Score 461; DB 4; Length 241;

Best Local Similarity 70.3%; Pred. No. 5e-42;

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 2 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61
 DB 51 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 110
 QY 62 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 111
 DB 111 PSNRETOOKSNLELRISLLIQSLWLPVQIGTPRFVNHLCGS-----HLVE 169
 QY 112 ALYVCG--ERGFYTPKRGIVEQ 134
 DB 170 GIOTLMGRLEDG---SPRTGQIFKQ 191

RESULT 11

US-09-280-030-66

Sequence 66; Application US/09280030A
 Patent No. 6506595
 GENERAL INFORMATION:
 APPLICANT: Sato, Seiji
 APPLICANT: Higashikuni, Naohiko
 APPLICANT: Kudo, Toshiyuki
 APPLICANT: Kondo, Masazaki
 TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
 TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
 FILE REFERENCE: 382.1026
 CURRENT APPLICATION NUMBER: US/09/280,030A
 CURRENT FILING DATE: 1999-03-26
 EARLIER APPLICATION NUMBER: JP10-87339/1998
 EARLIER FILING DATE: 1998-03-31
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 66
 LENGTH: 245
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:


```

Matches 102: Conservative 7: Mismatches 20: Indels 11: Gaps 4:
QY 2 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 27 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 62 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 87 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
   : 111 111 111 111 111 111 111 111 111 111 111 111
DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 14
US-09-465-461-1:
: Sequence 1, Application US/09465461
: Patent No. 6348444
: GENERAL INFORMATION:
: APPLICANT: CHAPPEL, Scott
: TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and immune response
: FILE OF INVENTION: after hematopoietic stem cell transplantation in humans
: FILE REFERENCE: CHAPPEL-6.1
: CURRENT APPLICATION NUMBER: US/09/465,461
: PRIOR APPLICATION NUMBER: 63/112,668
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: Patent in version 3.1
: SEQ ID NO 1
: LENGTH: 191
: TYPE: PRT
: ORGANISM: homo sapiens
US-09-465-461-1

Query Match 57.1% Score 455: DB 1: Length 191:
Best Local Similarity 69.7% Pred No. 2e-41:
Matches 101: Conservative 7: Mismatches 21: Indels 18: Gaps 1:
QY 2 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 1 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 62 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 61 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
   : 111 111 111 111 111 111 111 111 111 111 111 111
DB 120 GIOTLMGRLEDG---SPRTGQIFKQ 167

```

```

RESULT 15
US-09-465-461-1:
: Sequence 4, Application US/08187756C
: Patent No. 6397769
: GENERAL INFORMATION:
: APPLICANT: ROSEN, ET AL.
: TITLE OF INVENTION: Human Growth Hormone
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CAHELIA, BYRN, BAYN, SULFILLAN,
: STREET: 6 BECKER FARM ROAD
: CITY: ROSELAND
: STATE: NEW JERSEY
: COUNTRY: USA
: ZIP: 07068
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 INCH DISKETTE
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS

```

```

SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/187,756C
FILING DATE: January 27, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1750
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-187-756C-4

Query Match 57.1% Score 455: DB 1: Length 217:
Best Local Similarity 69.7% Pred No. 2e-41:
Matches 101: Conservative 7: Mismatches 21: Indels 18: Gaps 4:
QY 2 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 27 FPTIPSLRLEFNAMLAHRLHQLAFDTYOEFEAYTPKEOKYSFIONPOTSLSFSSEIPT 61
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 62 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
DB 87 PSNREETOQKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HVE 111
   : 111 1111 1111 1111 1111 111 111 111 111 111 111 111
QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
   : 111 111 111 111 111 111 111 111 111 111 111 111
DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

Search completed: September 16, 2003, 12:41:25
CPU time : 24.93 secs

```

GenCore version: 5.1.1.6
Copyright: (c) 1993 - 2003 GenCore Ltd.

OM protein - protein search, using sw mode.

Run on: September 16, 2003, 12:38:55 : Search time 47.2753 seconds
(without alignments)
472.415 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPTPLSLKFDNAYLRAER.....IVGQCTSTCSLYALENVCN 150

Scoring table: BLASTUM62

Gapop 10.0 , Gapext 0.5

Searched: 556259 seqs, 148393369 residues

Total number of hits satisfying chosen parameters: 556259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PCB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PCB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PCB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PCB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PCB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PCBOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PCB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PCBOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PCBOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PCBOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PCBOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PCB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US09_PCBOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PCBOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PCBOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PCBOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_PCBOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_PCBOMB.pep.*

Pred. No. is the number of residues predicted by Clustal to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	797	100.0	150	14	US-10-054-873-7
2	555.5	69.7	107	14	US-10-054-873-6
3	470	59.0	92	14	US-10-054-873-2
4	461	57.8	191	11	US-09-984-019-23
5	461	57.8	191	12	US-10-153-207-1
6	461	57.8	191	12	US-10-403-377-1
7	461	57.8	191	12	US-10-456-708-1
8	461	57.8	191	12	US-10-238-148-1
9	461	57.8	214	12	US-10-153-207-6
10	461	57.8	217	9	US-09-929-918-9
11	461	57.6	245	9	US-09-280-035-66
12	455	57.1	217	9	US-09-853-689-2
13	455	57.1	217	12	US-09-969-7482-4
14	454	57.0	217	10	US-09-804-409A-16
15	448	56.2	217	5	US-09-853-688-4

445 55.4 291 12 US-09-824-200-12 Sequence 12, Appl
440 55.2 191 12 US-10-306-822-1 Sequence 1, Appl
433 54.3 191 12 US-10-289-845-3 Sequence 3, Appl
399 50.1 217 9 US-09-850-887-3 Sequence 3, Appl
381 47.8 163 15 US-10-043-487-350 Sequence 350, Appl
381 47.8 191 12 US-10-153-207-2 Sequence 2, Appl
373 46.4 243 15 US-10-103-313-411 Sequence 411, Appl
345 43.3 246 15 US-10-188-246-18 Sequence 18, Appl
345 43.3 190 12 US-10-153-207-3 Sequence 3, Appl
331.5 37.8 182 10 US-09-861-687-19 Sequence 19, Appl
294 37.0 52 14 US-10-054-873-5 Sequence 5, Appl
286 35.9 147 9 US-09-736-611-8 Sequence 8, Appl
286 35.9 147 9 US-09-740-359-7 Sequence 7, Appl
284.5 35.7 124 10 US-09-894-711-18 Sequence 18, Appl
284.5 35.7 144 9 US-09-736-611-6 Sequence 6, Appl
284.5 35.7 144 9 US-09-740-359-5 Sequence 5, Appl
284.5 35.7 146 10 US-09-894-711-5 Sequence 5, Appl
278.5 34.9 51 11 US-09-858-935B-5 Sequence 5, Appl
278.5 34.9 51 14 US-10-028-410-3 Sequence 3, Appl
278 34.9 117 9 US-09-280-030-63 Sequence 63, Appl
277 34.8 96 10 US-09-947-563-4 Sequence 4, Appl
275.5 34.6 124 9 US-09-736-611-12 Sequence 12, Appl
275.5 34.6 124 9 US-09-740-359-12 Sequence 12, Appl
275.5 34.6 124 10 US-09-894-711-12 Sequence 12, Appl
275.5 34.6 125 9 US-09-736-611-10 Sequence 10, Appl
275.5 34.6 125 10 US-09-740-359-10 Sequence 10, Appl
273 34.5 50 14 US-10-056-009A-3 Sequence 3, Appl
271 34.0 96 10 US-09-947-563-5 Sequence 5, Appl
270 33.9 110 9 US-09-205-658-125 Sequence 125, App

ALIGNMENTS

RESULT: 1
US-10-054-873-7
: Sequence 7, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gen. Zhong Hu
: TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chapterone-Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/0/054.873
: FILING DATE: 22-Jag-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WC PCT/CN98/00052
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 11-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 45,946
: REFERENCE/DOCKET NUMBER: 020167-000130US
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 amino acids
: TYPE: amino acid

```

1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP----- 49
2  STRANDEDNESS: <Unknown>
3  TOPOLOGY: linear
4  MOLECULE TYPE: protein
5  SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match: 100.0%; Score 797; LP 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
DB 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
QY 61  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
DB 1  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
QY 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150
DB 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150

RESULT 2
US-10-054-873-6
Sequence 6, Application US/16054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong R.
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: P2-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-054-873-6

Query Match: 69.7%; Score 555.5; DB 14; Length 107;
Best Local Similarity 71.3%; Pred. No. 5.9e-54;
Matches 107; Conservative 0; Mismatches 0; Indels 4; Gaps 10

QY 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
DB 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
QY 61  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
DB 61  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
QY 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150
DB 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150

RESULT 4
US-09-964-010-23
Sequence 23, Application US/09984010
Publication No. US20030104578A1
GENERAL INFORMATION:

```

```

1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP----- 49
2  STRANDEDNESS: <Unknown>
3  TOPOLOGY: linear
4  MOLECULE TYPE: protein
5  SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match: 100.0%; Score 797; LP 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
DB 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
QY 61  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
DB 1  TPNREETQOKSNLELRISLLIQLSKWLPVGLTGPSPFVNOHLCGSHLVEALYLVCGER 120
QY 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150
DB 121  GFYTPKTRGIVEQCCISCSLYQLENYCN 150

RESULT 2
US-10-054-873-2
Sequence 2, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong R.
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: P2-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-054-873-2

Query Match: 59.0%; Score 470; DB 14; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.5e-44;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
DB 1  MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNPQTSLSFSESIP 60
QY 61  TPNREETQOKSNLELRISLLIQLSKWLPVQ 92
DB 61  TPNREETQOKSNLELRISLLIQLSKWLPVQ 92

RESULT 4
US-09-964-010-23
Sequence 23, Application US/09984010
Publication No. US20030104578A1
GENERAL INFORMATION:

```

```

1 APPLICANT: Ealience, David James
2 TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
3 AND SERUM ALBUMIN
4 NUMBER OF SEQUENCES: 2
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: FINNPFM, HENDERSON, PARABOW, GARRETT & DENNER, LLP
7 STREET: 1300 I Street, NW
8 CITY: Washington
9 STATE: DC
10 COUNTRY: USA
11 ZIP: 20005-3815
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC Compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/09/984,010
19 FILING DATE: 21-May-2002
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US 09/091,873
22 FILING DATE: 25-JUN-1998
23 APPLICATION NUMBER: PCT/SB96/03164
24 FILING DATE: 19-DEC-1996
25 INFORMATION FOR SEQ ID NO: 23:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 191 amino acids
28 TYPE: amino acid
29 STRANDEDNESS: 
30 TOPOLOGY: linear
31 MOLECULE TYPE: protein
32 HYPOTHEetical: NO
33 ANTI-SENSE: NO
34 SEQUENCE DESCRIPTION: SEQ ID NO: 23:
35 US-09-984-010-23
36
37 Query Match 57.8%, Score 461, DR 117 Length 191
38 Best Local Similarity 70.3%, Pred. No. 3,90-43
39 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4
40
41 QY 2 FTPLSLLENAMLAHRLHQAFDYQHEEAYIPKOKYSFLONPOTSIPTSPESIPT 61
42 IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
43 Db 1 FTTPLSLEFNAMLAHRLHQAFDYQHEEAYIPKOKYSFLONPOTSIPTSPESIPT 60
44 IFTTPLSLEFNAMLAHRLHQAFDYQHEEAYIPKOKYSFLONPOTSIPTSPESIPT 60
45 QY 62 PENRETQCKSNIELRSTLLICSMETPEVLTGTFPVNCHNGS-----ELVE 111
46 IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
47 Db 61 PSKRRETQCKSNIELRSTLLICSMETPEVLTGTFPVNCHNGS-----ELVE 111
48 IISKRRETQCKSNIELRSTLLICSMETPEVLTGTFPVNCHNGS-----ELVE 111
49 QY 112 AYLVCG--ERFFYPKTRGLIVP 134
50 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
51 Db 120 GQTIMGSEITG---SPRTGSIRPG 141
52 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
53
54 RESULT 5
55 US-10-153-207-1
56 : Sequence 1, Application: US/10153207
57 : Publication No. US20030153093A1
58 : GENERAL INFORMATION:
59 : APPLICANT: James A. Wells
60 : APPLICANT: Brian C. Cunningham
61 : TITLE OF INVENTION: GROWTH HORMONE VARIANTS
62 : FILE REFERENCE: 669.12-US-C7
63 : CURRENT APPLICATION NUMBER: US/10-153,207
64 : CURRENT FILING DATE: 2002-05-22
65 : PRIOR APPLICATION NUMBER: 08/479,864
66 : PRIOR FILING DATE: 1995-06-07
67 : PRIOR APPLICATION NUMBER: 08/190,723
68 : PRIOR FILING DATE: 1994-02-02
69 : PRIOR APPLICATION NUMBER: 07/960,227
70 : PRIOR FILING DATE: 1992-10-13
71 : PRIOR APPLICATION NUMBER: 07/7875,204
72 : PRIOR FILING DATE: 1992-04-27
73 : PRIOR APPLICATION NUMBER: 07/428,066

```



```

: APPLICANT: Cox III, George N
: TITLE OF INVENTION: Holder Biotechnology, Inc.
: FILE REFERENCE: 4152-1-PUS
: CURRENT APPLICATION NUMBER: US/12/400,728
: CURRENT FILING DATE: 2003-03-26
: PRIOR APPLICATION NUMBER: US/09/462,941
: PRIOR FILING DATE: 2002-01-14
: PRIOR APPLICATION NUMBER: 60/052,516
: PRIOR FILING DATE: 1997-07-14
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-400-708-1

Query Match      57.8%   Score 461   DB 12: Length 191
Best Local Similarity 70.3%   Pred. No. 3,90-43
Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 60
DB 1 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 60
QY 62 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
DB 61 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
QY 112 ALYVCG--ERGFFYTPKTRGIVEQ 134
DB 120 GQTLNGRLDGG---SPRTGQIFKO 141

```

```

RESULT 8
US-10-298-148-1
: Sequence 1, Application US/16258148
: Publication No. US20030171284A1
: GENERAL INFORMATION:
: APPLICANT: Cox III, George N
: TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
: FILE REFERENCE: 4152-1-PUS
: CURRENT APPLICATION NUMBER: US/10/298,148
: CURRENT FILING DATE: 2002-11-15
: PRIOR APPLICATION NUMBER: US/09/462,941
: PRIOR FILING DATE: 2003-01-14
: PRIOR APPLICATION NUMBER: 60/052,516
: PRIOR FILING DATE: 1997-07-14
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-298-148-1

Query Match      57.6%   Score 461   DB 12: Length 191
Best Local Similarity 70.3%   Pred. No. 3,90-43
Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 60
DB 1 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 60
QY 62 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
DB 61 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
QY 112 ALYVCG--ERGFFYTPKTRGIVEQ 134
DB 120 GQTLNGRLDGG---SPRTGQIFKO 141

```

```

RESULT 9
US-10-153-207-6
: Sequence 6, Application US/10153207
: Publication No. US20030153003A1
: GENERAL INFORMATION:
: APPLICANT: James A. Wells
: APPLICANT: Brian C. Cunningham
: TITLE OF INVENTION: GROWTH HORMONE VARIANTS
: FILE REFERENCE: 569-12-US-C7
: CURRENT APPLICATION NUMBER: US/10/153,207
: CURRENT FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: 08/479,884
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 05/190,723
: PRIOR FILING DATE: 1994-02-02
: PRIOR APPLICATION NUMBER: 07/960,227
: PRIOR FILING DATE: 1992-10-13
: PRIOR APPLICATION NUMBER: 07/875,204
: PRIOR FILING DATE: 1992-04-27
: PRIOR APPLICATION NUMBER: 07/428,066
: PRIOR FILING DATE: 1989-10-26
: PRIOR APPLICATION NUMBER: 07/264,611
: PRIOR FILING DATE: 1988-10-28
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 214
: TYPE: PRT
: ORGANISM: Homo Sapiens
US-10-153-207-6

Query Match      57.8%   Score 461   DB 12: Length 214
Best Local Similarity 70.3%   Pred. No. 4,5e-43
Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4

QY 2 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 61
DB 24 FPTPLSRFDNMLRAHCHQLADFYQEFEEAYIPKREKYSFQNPOTSLSPESIPT 83
QY 62 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
DB 84 PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCCS-----HLVE 111
QY 112 ALYVCG--ERGFFYTPKTRGIVEQ 134
DB 143 GQTLNGRLDGG---SPRTGQIFKO 164

RESULT 10
US-09-929-918-9
: Sequence 9, Application US/09929918
: Patent No. US20020090678A1
: GENERAL INFORMATION:
: APPLICANT: Kordyum, Vitaliy A.
: APPLICANT: Cherdynko, Svetlana I.
: APPLICANT: Slavchenko, Iryna Yu.
: APPLICANT: Vozianov, Oleksandr
: TITLE OF INVENTION: PHAGE-DEPENDENT SUPER PRODUCTION OF BIOLOGICALLY ACTIVE PROTEIN AND PEPTIDES
: FILE REFERENCE: PHAGE.006A
: CURRENT APPLICATION NUMBER: US/09/929,918
: CURRENT FILING DATE: 2001-08-15
: PRIOR APPLICATION NUMBER: 09/318,288
: PRIOR FILING DATE: 1999-05-25
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 217
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-929-918-9

```


GenCode version 5.1.6
Copyright (c) 1993 - 2003 Comptel, Ltd.

OM protein - protein search, using sw mode:

Run on: September 16, 2003, 12:36:15 : Search time 17.5987 seconds
(without alignments)
823,845 Million cell updates/sec

Title: US-09-423-100-7
Perfect score: 797
Sequence: 1 MFPIPLSRLEFDMALRAHR.....IVEUCCLSIQSLVLENYEN 150

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 56158632 residues
Total number of hits satisfying chosen parameters: 283305

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 43 summaries

Database : PIR_76:★
1: pir1:★
2: pir2:★
3: pir3:★
4: pir4:★

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	57.8	217	1	STHCV
2	460	57.7	217	2	somatotropin - pre
3	426.5	53.5	217	1	STHUV
4	426.5	53.5	256	1	somatotropin - pre
5	407.5	51.1	217	2	STHUV2
6	405	50.8	217	2	chorionic somatopa
7	396	49.7	212	2	somatotropin - the
8	396	49.7	217	2	chorionic somatopa
9	381	47.8	217	1	chorionic somatopa
10	381	47.8	217	2	chorionic somatopa
11	359.5	45.1	215	2	chorionic somatopa
12	310.5	39.0	216	2	chorionic somatopa
13	307.5	38.6	190	2	somatotropin - pre
14	306.5	38.5	190	1	somatotropin - pre
15	304.5	38.2	215	1	STHVS
16	302.5	38.0	216	1	STHVS
17	302.5	38.0	216	2	STHVS
18	301.5	37.8	190	2	STHVS
19	301.5	37.8	216	1	STHVS
20	301.5	37.8	216	2	STHVS
21	301.5	37.8	216	2	STHVS
22	299.5	37.6	216	2	STHVS
23	297.5	37.3	190	1	STHVS
24	295.5	37.1	190	2	STHVS
25	289.5	36.3	217	1	STHVS
26	289.5	36.3	217	1	STHVS
27	289.5	36.3	217	1	STHVS
28	289.5	36.3	217	2	STHVS
29	278.5	34.9	216	2	STHVS

30 277.5 34.8 110 1 INRB
31 277.5 34.8 110 2 B42179
32 275.5 34.6 226 2 A50509
33 275 34.5 96 2 PC7082
34 273.5 34.3 51 1 INWRP
35 273.5 34.3 51 1 INWRP
36 273.5 34.3 51 1 INEC
37 273.5 34.3 110 2 J00176
38 272 34.1 110 2 A42179
39 271.5 34.1 51 1 INHY
40 270 33.9 110 1 IFHU
41 268.5 33.7 51 1 INMSSP
42 268.5 33.7 131 2 A60625
43 267.5 33.6 51 2 A59151
44 266.5 33.4 105 1 IFBO
45 265.5 33.3 110 2 148166

ALIGNMENTS

RESULT 1
STHCV

somatotropin 1 precursor [validated] - human
N:Alternate names: growth hormone 1; hGH-N; pituitary somatotropin
N:Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, s
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 10-Feb-1995 #text_change 08-Dec-2000
C:Accession: A93731; A32435; A93694; A94247; A90051; A93778; A91764; A90217
R:DeNoto, F.M.; Moore, D.D.; Goodman, H.M.
Nucleic Acids Res. 9, 3719-3730, 1981
A:Title: Human growth hormone DNA sequence and mRNA structure: possible alternative
A:Reference number: A93731; MUID:82014939; PMID:6269091
A:Accession: A93731
A:Molecule type: DNA
A:Residues: 1-217 <DEN>
A:Cross-references: GH:V00520
A:Note: The 20K short form somatotropin lacks residues 58-72 (32-46 in the active h
R:Chou, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seebur
Genomics 4, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A:Reference number: A32435; MUID:89307277; PMID:2744760
A:Accession: A32435
A:Molecule type: DNA
A:Residues: 1-217 <CHE>
A:Cross-references: GH:J03071; NID:9183148; PIDN:AAAS2549.1; PID:q183149
R:Koska, W.; Rougeon, F.
Nucleic Acids Res. 7, 305-320, 1979
A:Title: Molecular cloning and nucleotide sequence of the human growth hormone struc
A:Reference number: A93694; MUID:80034477; PMID:385281
A:Accession: A93694
A:Molecule type: mRNA
A:Residues: 1-217 <ROS>
A:Cross-references: GH:V00519
A:Note: 35-550 was also found
R:Marital, J.A.; Halliwell, R.A.; Baxter, J.D.; Goodman, H.M.
Science 205, 602-607, 1979
A:Title: Human growth hormone: complementary DNA cloning and expression in bacteria
A:Reference number: A94247; MUID:79203293; PMID:377496
A:Accession: A94247
A:Molecule type: mRNA
A:Residues: 1-217 <MAR>
R:Li, C.H.; Dixon, J.S.; Liu, W.K.
Arch. Biochem. Biophys. 133, 70-91, 1969
A:Title: Human pituitary growth hormone. XIX. The primary structure of the hormone.
A:Reference number: A90046; MUID:59289202; PMID:5810834
A:Contents: annotation
R:Li, C.H.; Dixon, J.S.
Arch. Biochem. Biophys. 146, 233-236, 1971
A:Title: Human pituitary growth hormone. XXII. The primary structure of the hormone
A:Reference number: A90051; MUID:72143935; PMID:5144027
A:Accession: A90051
A:Molecule type: protein

Ds 87 PSNREETQOKSN: ELLR:SLLLIQSWLEPVQL 117

RESULT 3

STRUV
somatotropin 2 precursor - human
N:Alternate names: growth hormone variant-2; placental somatotropin form 2
C:Species: Homo sapiens (man)
C:Date: 17-Dec-1982 #sequence_revision 10-Feb-1995 #text_change 21-Jul-1995
C:Accession: D32435; H28072; A31511; 152163; A60711
R:Chen, F.Y.; Zhao, Y.C.; Smith, D.E.; Barrera Saldaña, H.A.; Bellizzi, R.M.; Genomics 4, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, cloning, and evolution.
A:Reference number: A32435; MUID:8307277; PMID:2743768
A:Accession: D32435
A:Molecule type: DNA
A:Residues: 1-217 <CHE>
A:Cross-references: G3:J03071; NID:8183148; PIDN:AAA52552.1; PIDN:ELF152
R:Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhaver, S.A.
J. Biol. Chem. 263, 9001-9006, 1988
A:Title: Two distinct species of human growth hormone-variant mRNA in the human placenta
A:Reference number: A92725; MUID 86243769; PMID:3379057
A:Accession: B28072
A:Molecule type: mRNA
A:Residues: 1-217 <COG>
R:Seeburg, P.H.
DNA 1, 239-249, 1982
A:Title: The human growth hormone gene family: nucleotide sequences show recent divergence
A:Reference number: A61511; MUID:83182010; PMID:7165039
A:Accession: A61511
A:Molecule type: DNA
A:Residues: 1-347; 36-217 <SEP>
R:Ignotz, A.; Scippo, M.L.; Frankenne, F.; Hennan, G.
Arch. Int. Physiol. Biochim. 96, 63-67, 1988
A:Title: Cloning and nucleotide sequence of placental hGH-V cDNA.
A:Reference number: 152164; MUID:89024984; PMID:2460050
A:Accession: 152164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <USO>
A:Cross-references: GB:M84851; NID:8183148; PIDN:AAA55551.1; PIDN:ELF152
R:Frankenne, F.; Scippo, M.L.; Van Boesum, J.; Ignotz, A.; Hennan, G.
J. Clin. Endocrinol. Metab. 71, 1518, 1990
A:Title: Identification of placental human growth hormone as the at with hormone V gene
A:Reference number: A60711; MUID:8317018; PMID:2166288
A:Accession: A60711
A:Molecule type: protein
A:Residues: 27-44; 46-57 <RA>
A:Experimental source: tissue placenta
A:Note: partial glycosylation was demonstrated by protein binding
C:Comment: This gene is expressed by the placenta.
C:Genetics:
A:Gene: GDB:GR2
A:Cross-references: GDB:119983; OMIM:139240
A:Map position: 17q22-17q24
A:Introns: 4/1: 57/3; 97/3; 152/3
C:Superfamily: prolactin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-217/product: somatotropin 2, long splice form #status predicted <SIG>
F:79-191,208-215/disulfide bonds: #status predicted
F:166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5% Score 426.5; DB 1; Length 217;
Best Local Similarity 78.1% Pred. No. 1.8e-34;
Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

Ds 2 FPTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 61
|||||
Ds 27 FPTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 96

OY 62 PSNREETQOKSN:ELRLISLLIQSWLEPVQL-----GTGPRFVNQRL 104
|||
Ds 87 PSNREETQOKSN:ELRLISLLIQSWLEPVQLLRVSFANSLVYGASDSNVYRHL 140
|||

RESULT 4

STRUV
somatotropin 2 precursor, splice form 2 - human
N:Alternate names: growth hormone variant-2; placental somatotropin form 2
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1989 #sequence_revision 10-Feb-1995 #text_change 02-Sep-1997
C:Accession: A28072
R:Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhaver, S.A.
J. Biol. Chem. 263, 9001-9006, 1988
A:Title: Two distinct species of human growth hormone-variant mRNA in the human placenta
A:Reference number: A92725; MUID:86243769; PMID:3379057
A:Accession: A28072
A:Molecule type: mRNA
A:Residues: 1-256 <COG>
A:Note: an alternative splice junction for intron 4 is used
C:Genetics:
A:Gene: GDB:GH2
A:Cross-references: GDB:119983; OMIM:139240
A:Map position: 17q22-17q24
A:Introns: 4/1: 57/3; 97/3; 152/3
C:Superfamily: prolactin
C:Keywords: alternative splicing; hormone; placenta
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-256/product: somatotropin 2 splice form 2 #status predicted <MAT>
Query Match 53.5% Score 426.5; DB 1; Length 256;
Best Local Similarity 78.1% Pred. No. 2.2e-34;
Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;
OY 2 FPTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 61
Ds 27 FPTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 86
OY 62 PSNREETQOKSN:ELRLISLLIQSWLEPVQL-----GTGPRFVNQRL 104
|||||
Ds 87 PSNREETQOKSN:ELRLISLLIQSWLEPVQLLRVSFANSLVYGASDSNVYRHL 140
|||||

Query Match 51.1% Score 407.5; DB 2; Length 217;
Best Local Similarity 71.9% Pred. No. 1.4e-32;
Matches 82; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

OY 3 FTLPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 62
|||||
Ds 28 PSVPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLONPOTSLSFSESIPT 87
|||||

OY 63 SNREETQOKSN:ELRLISLLIQSWLEPVQLTGPRFVNQRLCGSHLVYALYLV 116
|||||
Ds 88 SNREETQOKSN:ELRLISLLIQSWLEPVQL-LGSVFANSLVYGTSSESDAYOLL 140
|||||

```

RESULT 6
167411
somatotropin - rhesus macaque
N:Alternate names: growth hormone
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I67411
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.F.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I67411
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:L16553; NID:q293110; PID:AAA18839.1; PID:q293109
C:Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 217;
Best Local Similarity 82.2%; Pred. No. 1.8e-31;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PTPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 62
DQ 23 PSVPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 82
QY 63 SNRETOOKSNLELRISLLIQLSWLEPVQ 92
DQ 83 SNLEETOOKSNLELRISLLIQLSWLEPVQ 112

RESULT 7
167408
chorionic somatomammotropin-2 - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I67408
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.F.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I67408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <RES>
A:Cross-references: GB:L16553; NID:q293110; PID:AAA18840.1; PID:q293111
C:Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 212;
Best Local Similarity 82.2%; Pred. No. 1.8e-31;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PTPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 62
DQ 23 PSVPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 82
QY 63 SNRETOOKSNLELRISLLIQLSWLEPVQ 92
DQ 83 SNLEETOOKSNLELRISLLIQLSWLEPVQ 112

RESULT 8
153267
chorionic somatomammotropin-1 - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I53267
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.F.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I53267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:L16552; NID:q293108; PID:AAA18839.1; PID:q293109
C:Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 217;
Best Local Similarity 82.2%; Pred. No. 1.8e-31;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PTPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 62
DQ 23 PSVPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSESISPTP 82
QY 63 SNRETOOKSNLELRISLLIQLSWLEPVQ 92
DQ 83 SNLEETOOKSNLELRISLLIQLSWLEPVQ 117

RESULT 9
153267
chorionic somatomammotropin A precursor [validated] - human
R:Alternate names: chorionic somatomammotropin; placental lactogen
C:Species: Homo sapiens (man)
C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 08-Dec-2000
C:Accession: C32435; #944422; 152342; A93833; A93192; A90054; A94427; A61283; 155229;
R:Chen, E.Y.; Liang, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg
Endocrinology 47, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A:Reference number: A94422; MID:89307277; PMID:2744760
A:Accession: C32435
A:Molecule type: DNA
A:Residues: 1-217 <CHR>
A:Cross-references: GB:C03671; NID:q183148; PID:AAA52551.1; PID:q183151
R:Goodman, H.M.; Denoto, F.; Fiddes, J.C.; Hallewell, R.A.; Page, G.S.; Smith, S.;
In Mobilization and Reassembly of Genetic Information. Scott, W.A.; Werner, R.; Jose
A:Reference number: A94422
A:Accession: A94422
A:Molecule type: mRNA
A:Residues: 1-217 <GOO>
R:Tanaka, K.; Masuda, N.; Watahiki, M.; Yamakawa, M.; Shimizu, K.; Nagai, J.; Nakast
Biochem. Int. 16, 287-292, 1988
A:Title: cDNA cloning of human chorionic somatomammotropin-1 mRNA whose transcriptic
A:Reference number: 152342; MID:88209056; PMID:2835050
A:Accession: I52342
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3 <TAN>
A:Cross-references: GB:M35419; NID:q506822
R:Sherwood, L.M.; Burstein, Y.; Schechter, I.
Proc. Natl. Acad. Sci. U.S.A. 76, 3819-3823, 1979
A:Title: Primary structure of the NH-2-terminal extra piece of the precursor to huma
A:Reference number: A93833; MID:80034970; PMID:291043
A:Accession: A93833
A:Molecule type: protein
A:Residues: 1-3-26 <SHE>
A:Experimental source: placenta
R:Shine, J.; Seeburg, P.H.; Martini, J.A.; Baxter, J.D.; Goodman, H.M.
Nature 270, 494-499, 1977
A:Title: Construction and analysis of recombinant DNA for human chorionic somatomamm
A:Reference number: A93192; MID:78071761; PMID:593368
A:Accession: A93192
A:Molecule type: DNA
A:Residues: 50-217 <SHI>
A:Experimental source: placenta
R:Li, C.H.; Dixon, J.S.; Chung, D.
Arch. Biochem. Biophys. 155, 95-110, 1973
A:Title: Amino acid sequence of human chorionic somatomammotropin.
A:Reference number: A90054; MID:73201971; PMID:4712450
A:Accession: A90054

```

A:Molecule type: protein
 A:Residues: 27-217 <LJC>
 A:Experimental source: placenta
 R:Niall, H.D.
 in Prolactin and Carcinogenesis. Proc. Fourth Toronto Workshop: Prolactin, Growth, and
 A:Title: The chemistry of the human lactogenic hormones.
 A:Reference number: A94427
 A:Accession: A94427
 A:Molecule type: protein
 A:Residues: 27-217 <NIA>
 A:Experimental source: placenta
 R:Nic A Baird, N.; Tipton, K.F.
 Biochem. Soc. Trans. 19, 20S, 1971
 A:Title: Catechol-O-methyltransferase from human placenta: purification and some properties
 A:Reference number: A61283; MUID:9124406; PMID:2037146
 A:Accession: A61283
 A:Molecule type: protein
 A:Residues: 27-46 <NIC>
 A:Note: Chorionmammotropin apparently copurified with placental catechol-O-methyltransferase
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Lauer, M.
 Nature New Biol. 233, 59-61, 1971
 A:Title: Amino-acid sequence of human placental lactogen.
 A:Reference number: A93403; MUID:72016313; PMID:5286383
 A:Contents: annotation
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Lauer, M.
 Nature New Biol. 235, 64, 1972
 A:Reference number: A93405
 A:Contents: annotation
 R:Schneider, A.B.; Kovalski, K.; Russell, J.; Sherwood, L.M.
 J. Biol. Chem. 254, 3782-3787, 1979
 A:Title: Identification of the interchain disulfide bonds of dimeric human placental lac
 A:Reference number: A92251; MUID:79173081; PMID:438159
 A:Contents: annotation; dimeric disulfide bonds
 R:Seiby, M.; Barta, A.; Baxter, J.D.; Beil, G.J.; Eberhardt, N.L.
 J. Biol. Chem. 259, 13332-13338, 1984
 A:Title: Analysis of a major human chorionic somatomammotropin gene. Evidence for two in
 A:Reference number: 155229; MUID:85030426; PMID:6239192
 A:Accession: 155229
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:X02401; NID:q181120; PIDN:AAA521151; PID:q181121
 R:Seeburg, P.H.; Shine, J.; Martial, J.A.; Ullrich, A.; Goodman, H.
 Trans. Assoc. Am. Physicians 90, 108-116, 1977
 A:Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.
 A:Reference number: 159658; MUID:78160787; PMID:611657
 A:Accession: 159658
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: tRNA
 A:Residues: 160-217 <RE2>
 A:Cross-references: GB:M25118; NID:q181124; PIDN:AAA521151; PID:q181125
 C:Genetics:
 A:Gene: GDB:CSH1
 A:Cross-references: GDB:119084; OMIM:155200
 A:Map position: 17q22-17q24
 A:Introns: 4/1; 57/3; 97/3; 152/1
 C:Superfamily: prolactin
 C:Keywords: hormone; placenta
 F:1-26/Domain: signal sequence #status experimental <SIG>
 F:27-217/Product: chorionmammotropin A #status experimental <MAT>
 F:79-191/Disulfide bonds: #status experimental
 F:208-215/Disulfide bonds: (in monomeric form) #status experimental
 F:208/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental
 F:215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental
 Query Match 47.8%; Score 381; DB 2; Length 217;
 Best Local Similarity 82.0%; Pred. No. 5.4e-30;
 Matches 73; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 QY 4 TIPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 63
 DB 29 TVPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 63

QY 64 NREETOQKSNLELRISLLLIQSWLEPVQ 92
 DB 89 NMEETOQKSNLELRISLLLIQSWLEPV 117
 RESULT 10
 Chorionmammotropin B precursor - human
 R:2445
 A:Alternate names: chorionic somatomammotropin; 2
 C:Species: Homo sapiens (man)
 C>Date: 29-Dec-1989 #sequence_revision 29-Dec-1989 #text_change 16-Jul-1999
 C:Accession: E32435
 R:Chen, E.Y.; Zhao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg
 Genomics 4, 479-497, 1989
 A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
 A:Reference number: A32435; MUID:89307277; PMID:2744750
 A:Accession: E32435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-217 <CHE>
 A:Cross-references: GB:303071; NID:q183148; PIDN:AAA52553.1; PID:q183153
 C:Genetics:
 A:Gene: GDB:CSH2
 A:Cross-references: GDB:119813; OMIM:118820
 A:Map position: 17q22-17q24
 C:Superfamily: prolactin
 Query Match 47.8%; Score 381; DB 2; Length 217;
 Best Local Similarity 82.0%; Pred. No. 5.4e-30;
 Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
 QY 4 TIPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 63
 DB 29 TVPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 83
 QY 64 NREETOQKSNLELRISLLLIQSWLEPVQ 92
 DB 89 NMEETOQKSNLELRISLLLIQSWLEPV 117
 RESULT 11
 A26449
 Chorionmammotropin precursor (allele hCS-3) - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 28-Jul-1995
 C:Accession: A26449
 R:Hecht, H.J.; Kimelman, M.J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.
 DNA 5, 59-70, 1987
 A:Title: The human growth hormone gene locus: structure, evolution, and allelic var
 A:Reference number: A26449; MUID:67161235; PMID:3030680
 A:Accession: A26449
 A:Molecule type: DNA
 A:Residues: 1-215 <HIR>
 C:Superfamily: prolactin
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-215/Product: chorionmammotropin, hCS-3 allele #status predicted <MAT>
 Query Match 45.1%; Score 359.5; DB 2; Length 215;
 Best Local Similarity 80.5%; Pred. No. 7e-28;
 Matches 70; Conservative 8; Mismatches 8; Indels 1; Gaps 1;
 QY 4 TIPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 63
 DB 29 TVPLSRLEFDNMLRAHRLHOLAFDTYQFEAYIPKEQKYSFLONPOTLSFSSESIPTPS 88
 QY 64 NREETOQKSNLELRISLLLIQSWLEPV 90
 DB 89 NMEETOQKSNLELRISLLLIQSWLEPV 114
 RESULT 12
 349159
 somatomotropin - golden hamster

Search completed: September 16, 2003, 12:40:56
Job time : 18.5097 secs

GenCore version 5.0.6
Copyright (c) 1993 - 2003 GenPharm Inc.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:33:32 : Search time 11.573.2 seconds
(without alignment)
504.233 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPIPLSRLEFNAALRAHR.....IVKCCISICSLVLENYCN 50

Scoring table: BLSCM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	451	57.8	217	1	SOMA_HUMAN
2	451	57.8	217	1	SOMA_PANTR
3	432	57.7	217	1	SOMA_MACRU
4	437	54.8	217	1	SOMA_SAIGR
5	434.5	54.5	217	1	SOMA_PANTR
6	432	54.2	217	1	SOMA_CALJA
7	426.5	53.5	217	1	SOMA_HUMAN
8	399	50.1	217	1	SOMA_MACRU
9	381	47.8	217	1	PGL_HUMAN
10	310.5	39.0	216	1	SOMA_MESAU
11	307.5	38.6	190	1	SOMA_PALABO
12	306.5	38.5	216	1	SOMA_HORSE
13	306.5	38.5	217	1	SOMA_GALSE
14	306.5	38.5	217	1	SOMA_XYCY
15	304.5	38.2	216	1	SOMA_XYSE
16	302.5	38.0	216	1	SOMA_RABIT
17	302.5	37.8	216	1	SOMA_RAT
18	301.5	37.8	190	1	SOMA_COXAF
19	301.5	37.8	216	1	SOMA_CANPA
20	301.5	37.8	216	1	SOMA_FELCA
21	301.5	37.6	216	1	SOMA_FIG
22	299.5	37.6	216	1	SOMA_RSVI
23	297.5	37.3	190	1	SOMA_LAMPA
24	295.5	37.1	190	1	SOMA_VULVU
25	291.5	36.6	215	1	SOMA_MONDO
26	291.5	36.6	215	1	SOMA_TRIVU
27	289.5	36.3	217	1	SOMA_BOVIN
28	289.5	36.3	217	1	SOMA_CEREE
29	289.5	36.3	217	1	SOMA_SHEEP
30	282.5	35.4	217	1	SOMA_RIBRU
31	278.5	34.9	216	1	SOMA_MELGA
32	277.5	34.8	110	1	INS_CRAE
33	277.5	34.8	110	1	INS_RABIT

34 275.5 34.6 216 1 SOMA_CHICK
35 274.5 34.4 217 1 SOMA_STRCA
36 273.5 34.3 51 1 INS_BALPH
37 273.5 34.3 51 1 INS_ELENA
38 273.5 34.3 110 1 INS_MACFA
39 272.5 34.2 190 1 SOMA_CRONO
40 272.5 34.2 190 1 INS_PANTR
41 270 33.9 110 1 INS_HUMAN
42 268.5 33.7 51 1 INS_ACOCA
43 258.5 33.7 191 1 SOMA_CHEMY
44 256.5 33.4 105 1 INS_BOVIN
45 246.5 33.4 110 1 INS_SPETR

ALIGNMENTS

RESULT 1
SOMA_HUMAN STANDARD; PRT: 217 AA.
AC P01241: Q14405; Q18631; Q9HE21; Q9UMJ7; Q9UNL5;
BT 22-JUN-1986 (Rel. 0); Created
CT 01-MAR-1992 (Rel. 2); Last sequence update;
ET 15-SEP-2003 (Rel. 42; Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
GN GH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A. (ISOFORM 1).
RA MEDLINE=80034477; PubMed=386281;
RA Koskam W., Rougeon F.;
RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."
RL Nucleic Acids Res. 7:305-320(1979).
RN 2;
RS SEQUENCE FROM N.A. (ISOFORM 2).
RA MEDLINE=79203293; PubMed=377456;
RA Martini J.A., Halliwell R.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression in bacteria."
RN 3;
RS SEQUENCE FROM N.A. (ISOFORM 1); AND POSSIBLE ALTERNATIVE SPLICING.
RA MEDLINE=2014939; PubMed=6256091;
RA Denoto F.N., Moore D.B., Goodman H.M.;
RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."
RN 4;
RS Nucleic Acids Res. 9:3719-3730(1981).
RS SEQUENCE FROM N.A.
RA MEDLINE=83182010; PubMed=7169009;
RA Seeburg P.H.;
RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."
RN 5;
RS DNA 1:239-243(1982).
RS SEQUENCE FROM N.A.
RA MEDLINE=69307277; PubMed=2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
RA Gellinas R.E., Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and evolution."
RN 6;
RS Genomics 4:479-497(1989).
RS SEQUENCE FROM N.A. (ISOFORM 3).
RA TISSUE=Pituitary;
RA Gu J., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;
RT "A novel gene expressed in human pituitary."
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RA Lewis U.J., Borewald L.F., Lewis L.J.;
 RT "The 20,000-dalton variant of human growth hormone: location of the
 PT amino acid deletions";
 RL Biochem. Biophys. Res. Commun. 92:511-516(1980).
 RN [17]
 RA DEAMINATION OF GDN-163 AND ASN-178.
 RX MEDLINE=82052997; PubMed=7028740;
 RA Lewis U.J., Singh R.N., Bonewald L.F., Seavey B.K.;
 RT "Altered proteolytic cleavage of human growth hormone as a result of
 PT deamidation";
 RL J. Biol. Chem. 256:11645-11650(1981).
 RN [18]
 RA REVIEW.
 RX MEDLINE=99321812; PubMed=10393484;
 RA Baumann G.;
 RT "Growth hormone heterogeneity in human pituitary and plasma";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RN [19]
 RA 3D-STRUCTURE MODELING.
 RX MEDLINE=88190073; PubMed=3447173;
 RA Cohen F.E., Kuntz I.D.;
 RT "Prediction of the three-dimensional structure of human growth
 PT hormone";
 RL Proteins 2:162-166(1987).
 RN [20]
 RA X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=92196577; PubMed=1549776;
 RA de Vos A.M., Uitsch M., Kossiakoff A.A.;
 RT "Human growth hormone and extracellular domain of its receptor:
 PT crystal structure of the complex";
 RL Science 255:306-312(1992).
 RN [21]
 RA X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=95075452; PubMed=794244;
 RA Somers W., Uitsch M., de Vos A.M., Kossiakoff A.A.;
 RT "The X-ray structure of a growth hormone-prolactin receptor complex";
 RL Nature 372:476-481(1994).
 RN [22]
 RA X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97113023; PubMed=8943276;
 RA Pavlovsk A.G., Wisdawa A.;
 RT "The crystal-structure of wild-type growth-hormone at 2.5-A
 PT resolution";
 RL Protein Pept. Lett. 2:333-340(1995).
 RN [23]
 RA X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97113023; PubMed=8943276;
 RA Sundstrom M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,
 RT "Crystal structure of an antagonist mutant of human growth hormone,
 PT G220K, in complex with its receptor at 2.9-A resolution";
 RL J. Mol. Biol. 277:32197-32203(1996).
 RN [24]
 RA VARIANT CYS-105.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 PT of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RN [25]
 RA ERRATUM.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 PT of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RN [26]
 RA FUNCTION: Plays an important role in growth control. Its major
 PT role in stimulating body growth is to stimulate the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC

RA SEQUENCE FROM N.A. (ISOFORM 4).
 RT TISSUE=pituitary;
 RX MEDLINE=20402571; PubMed=10941946;
 RA Hu Y.-J., Huang C.-H., Li Y.-H., Jiang C.-L., Fu S., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mo Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-H.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 PT axis and full-length cDNA cloning";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN [8]
 RA SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=8637393; PubMed=3912261;
 RA Gray G.L., Baldridge J.S., McKee K.S., Heyesker H.L., Chang C.N.;
 RT "Periplasmic production of correctly processed human growth hormone in
 PT Escherichia coli: natural and bacterial signal sequences are
 RT interchangeable";
 RL Gene 99:247-254(1985).
 RN [9]
 RA SEQUENCE OF 27-217.
 RX MEDLINE=69289202; PubMed=6810834;
 RA Li C.H., Dixon J.S., Liu W.K.;
 RT "Human pituitary growth hormone. XIX. The primary structure of the
 PT hormone";
 RL Arch. Biochem. Biophys. 131:70-91(1969).
 RN [10]
 RA SEQUENCE OF 27-217, AND REVISIONS.
 RX MEDLINE=72143935; PubMed=5144027;
 RA Li C.H., Dixon J.S.;
 RT "Human pituitary growth hormone. 32. The primary structure of the
 PT hormone: revision";
 RL Arch. Biochem. Biophys. 146:233-236(1971).
 RN [11]
 RA REVISION.
 RX MEDLINE=73092028; PubMed=4675454;
 RA Bewley T.A., Dixon J.S., Li C.H.;
 RT "Sequence comparison of human pituitary growth hormone, human
 PT chorionic somatomammotropin, and ovine pituitary growth and
 RT lactogenic hormones";
 RL Int. J. Pept. Protein Res. 4:261-267(1972).
 RN [12]
 RA SEQUENCE OF 27-61 AND 102-124.
 RX MEDLINE=71139765; PubMed=5229045;
 RA Niall H.D.;
 RT "Revised primary structure for human growth hormone";
 RL Nature New Biol. 230:90-91(1971).
 RN [13]
 RA REVISIONS TO 119-120 AND 157-159.
 RX MEDLINE=71151958; PubMed=5279528;
 RA Niall H.D., Hogan M.L., Sauter R., Rosenblum I.V., Greenwood F.C.;
 RT "Sequences of pituitary and placental lactogenic and growth hormones:
 PT evolution from a primordial peptide by gene reduplication";
 RL Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
 RN [14]
 RA REVISION.
 RX MEDLINE=81117361; PubMed=7462247;
 RA Chapman G.E., Rogers K.M., Erttaint T., Bradshaw K.A., Bates G.C.,
 RA Turner C., Garry P.D., Crane-Robinson C.;
 RT "The 20,000 molecular weight variant of human growth hormone.
 PT Preparation and some physical and chemical properties";
 RL J. Biol. Chem. 256:2395-2401(1981).
 RN [16]
 RA SEQUENCE OF 46-80 (ISOFORM 2).
 RX MEDLINE=80130196; PubMed=7356479;

```

Query Match      57.8% Score 461; DB 1; Length 217;
Best Local Similarity 70.3%; Pred. No. 4; Gap 34;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 41
DE 2 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 41
QY 27 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 41
DE 27 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 41
QY 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
DE 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
QY 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
DE 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
QY 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
DE 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
QY 146 GTQLMGRLDGG---SPRTGQIFKQ 167
DE 146 GTQLMGRLDGG---SPRTGQIFKQ 167

RESULT 2
SOMA_PANTR
ID SOMA_PANTR STANDARD: PRT: 217 AA.
AC PS9756;
DT 28-FEB-2003 (Rel. 4); Created)
DI 28-FEB-2003 (Rel. 4); Last sequence update)
DE 28-FEB-2003 (Rel. 4); Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
GN GH1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP Revol A., Esquivel D., Santiago D., Barrera-Saidana H.;
RT "Independent duplication of the growth hormone gene in three Anthropoid lineages."
RL Submitted (APR-2001) to the EMBL/GenBank/DBS databases.
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues (by similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@sib-sib.ch).
CC EMBL: AF374232; AML72284.1;
CC InterPro: IPR001450; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PRINTS: PK00836; SOMATOTROPIN.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
CC Hormone; Pituitary; Signal.
CC SIGNAL 1 26
CC CHAIN 27 217 BY SIMILARITY.
CC DISULFID 79 191 SOMATOTROPIN.
CC DISULFID 208 215 BY SIMILARITY.
CC CONFLICT 100 100 E -> Q (IN REF. 2).
CC CONFLICT 179 179 N -> D (IN REF. 2).
CC SEQUENCE 217 AA; 24843 MW; FEA295DEDE5E18674 CRC64;

Query Match      57.8% Score 461; DB 1; Length 217;
Best Local Similarity 70.3%; Pred. No. 4; Gap 34;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 61

```

```

DE 27 FPTPLSLFDNMLRAHRLHOLAFDYQEFEEAYIPKEQKYSFLQNTQSLSESIPT 86
QY 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
DE 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
QY 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
DE 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGPRFVNQHLGSS-----HIVE 111
QY 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
DE 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
QY 146 GTQLMGRLDGG---SPRTGQIFKQ 167
DE 146 GTQLMGRLDGG---SPRTGQIFKQ 167

RESULT 3
SOMA_MACHU
ID SOMA_MACHU STANDARD: PRT: 217 AA.
AC P13093;
DT 01-OCT-1993 (Rel. 27; Created)
DI 01-OCT-1994 (Rel. 30; Last sequence update)
DE 28-FEB-2003 (Rel. 4); Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
GN GH1.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=94008724; PubMed=8404617;
RA Golos I.G., Durning K., Fisher J.M., Fowler P.D.;
RT "Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta."
RL Endocrinology 133:1744-1752(1993).
RN [2]
RP SEQUENCE OF 27-217.
RP MEDLINE=86129460; PubMed=3080959;
RA Li C.H., Chung D., Lalm H.W., Stein S.;
RT "The primary structure of monkey pituitary growth hormone."
RA Arch. Biochem. Biophys. 245:287-291(1986).
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@sib-sib.ch).
CC EMBL: L16556; AAA18842.1;
CC PIR: P02241; IAXI.
CC PIR: P02241; IAXI.
CC InterPro: IPR001450; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
CC Hormone; Pituitary; Signal.
CC SIGNAL 1 26
CC CHAIN 27 217 SOMATOTROPIN.
CC DISULFID 79 191 BY SIMILARITY.
CC DISULFID 208 215 BY SIMILARITY.
CC CONFLICT 100 100 E -> Q (IN REF. 2).
CC CONFLICT 179 179 N -> D (IN REF. 2).
CC SEQUENCE 217 AA; 24913 MW; 2C5180341EEC46D0 CRC64;

```

```

Query Match      57.7%  Score 460; DB 1; Length 217;
Best Local Similarity 96.9%  Pred. No. 5.8e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #1
DB 27 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #6
QY 62 PSNREETOOKSNLELLRISLLIQSWLEPVQ #2
DB 87 PSNREETOOKSNLELLRISLLIQSWLEPVQ #117

RESULT 4
ID SOMA_SAIRB STANDARD; PRT; 217 AA.
AC P56143.
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Samiri belivensis belivensis (Belivian squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Samiri
OC NCBI_TaxID=39432;
FN [1]
PP SEQUENCE FROM N.A.
RA MEDLINE=21265430; PubMed=2371582;
RX Liu J.C., Makova K.E., Adkins R.M., Li W.H.
RT "Episodic evolution of growth hormone in primates and emergence of the
RT species specificity of human growth hormone receptor."
NL Mol. Biol. Evol. 18:945-953(2001).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues (ay similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF139040; AA62287.1;
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal; BY SIMILARITY.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
SQ SEQUENCES 217 AA; 24654 MW; 9515289992C529F7 CRC64;

Query Match      54.8%  Score 437; DB 1; Length 217;
Best Local Similarity 91.3%  Pred. No. 1e-35;
Matches 84; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #1
DB 27 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #6
QY 62 PSNREETOOKSNLELLRISLLIQSWLEPVQ #3
DB 87 PSNREETOOKSNLELLRISLLIQSWLEPVQ #117

```

```

ID SOM2_PANTR STANDARD; PRT; 217 AA.
AC P58757.
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
DE hormone) (Growth hormone 2).
GN GH2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
OC NCBI_TaxID=9596;
OX [1]
KP SEQUENCE FROM N.A.
RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.
RT "Independent duplication of the growth hormone gene in three
RT Anthropoid lineages."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF374233; AAL7228.1;
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Placenta; Signal; Glycoprotein.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 GROWTH HORMONE VARIANT.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
SQ SEQUENCES 217 AA; 24990 MW; 1592A429075677DE CRC64;

Query Match      54.5%  Score 434.5; DB 1; Length 217;
Best Local Similarity 78.9%  Pred. No. 1.8e-35;
Matches 90; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #1
DB 27 FTTPLSRLEFDMAMLRHRLHQLAFDYQEFEEAYIPKEQKYSFLQNPQTSLSSESISPT #6
QY 62 PSNREETOOKSNLELLRISLLIQSWLEPVQ #104
DB 87 PSNREETOOKSNLELLRISLLIQSWLEPVQ #140

RESULT 6
ID SOMA_CALJA STANDARD; PRT; 217 AA.
AC Q9GMB3;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DR Somatotropin precursor (growth hormone)
GN GH1
OS Callithrix jacchus (Common marmoset)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix
CX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallis O.C., Wallis M.
RT "Cloning and characterization of a putative growth hormone encoding
RL gene from the marmoset (Callithrix jacchus)".
RL Submitted (AUG-2003) to the EMBL/GenBank/CCDB databases.
CC 1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues (By similarity).
CC 2- SUBCELLULAR LOCATION: Secreted.
CC 3- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, in
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch).
CC
DR EMBL: AC297563; CAC03481.1;
DR HSSP: P01241; 1A22.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PROSITE: PS00266; SOMATOTROPIN_1: 1.
DR PROSITE: PS00338; SOMATOTROPIN_2: 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
SQ SEQUENCE 217 AA: 24959 MR: E102:51A12CEA192 CRG04:
Query Match 54.2%; Score 132; DS 1; Length 217;
Best local Similarity 91.2%; Pred. No. 3,26-16;
Matches #3: Conservative 43; Indels 0; Gaps 0;
QY 2 FTPTLSRFDNMLRAHPLKLCADFTVDFEATIPKQKYSPLQNTQTSFESPT 41
Db 27 FTPTLSRGLNMLRAHPLKLCADFTVDFEATIPKQKYSPLQNTQTSFESPT 40
QY 62 PSNREETQCKSNLELRISLLELQSWLRFVQ 52
Db 87 PASKKSTQCKSNLELRISLLELQSWLRFVQ 117
RESULT 7
SOM2_HUMAN
SD SOM2_HUMAN STANDARD; iRT: 217 AA.
AC P01242; PQ05P7;
DT 21-JUL-1986 (Rel. 01, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
ET 15-SEP-2003 (Rel. 42, Last annotation update)
DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
DE hormone) (Growth hormone 2).
GN GH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).

```

```

FX MEDLINE-83182010; PubMed-7169009;
FA Seeburg P.H.;
AT "The human growth hormone gene family: nucleotide sequences show
AT recent divergence and predict a new polypeptide hormone.";
FA DNA 1:239-249(1982).
FA [2]
FX SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
FX MEDLINE-8243769; PubMed-4379597;
RA Cooke N.E., Ray J., Emery C.G., Liebhafner S.A.;
PT "Two distinct species of human growth hormone-variant mRNA in the
PT human placenta predict the expression of novel growth hormone
PT proteins";
PL J Biol. Chem. 263:9001-9006(1988).
RX [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE-89021584; PubMed-2460050;
RA Iqbal A., Scippo M.L., Frankenne F., Hennen G.;
RT "Cloning and nucleotide sequence of placental hGH-V cDNA.";
FA Arch. Int. Physiol. Biochim. 96:63-67(1988).
FA [4]
RP SEQUENCE FROM N.A.
FX MEDLINE-89307277; PubMed-2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
RT evolution.";
RL Genomics 4:479-497(1989).
FX [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
FX MEDLINE-22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore L., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gutarane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Viatralo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M., Skalska U., Smailus D.E.,
RA Schurich A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
FA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
FX [6]
RX REVIEW
FX MEDLINE-93321812; PubMed-10393484;
FA Baumann G.;
RT "Growth hormone heterogeneity in human pituitary and plasma.";
FA Horm. Res. 51 Suppl. 1:2-6(1999).
CC 1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC 2- SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer, disulfide-
CC linked or non-covalently associated, in homopolymeric and
CC heteropolymeric combinations. Can also form a complex either with
CC GHBP or with the alpha2-macroglobulin complex.
CC 3- SUBCELLULAR LOCATION: Secreted.
CC 4- ALTERNATIVE PRODUCTS:
CC Name=1; Synonyms=GH-V1;
CC IsoId=P01242-1; Sequence=Displayed;
CC Name=2; Synonyms=GH-V2;

```


DT 01-APR-1988 (Rel. 37, last sequence update)
 DI 15-SEP-2003 (Rel. 42, last annotation update)
 DE Lactogen precursor (Chorionamniototropin) (Chorionic somatomammotropin).
 GN CSH1 AND CSH2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (GENE CSH1).
 RX MEDLINE=85033426; PubMed=6208192;
 RA Selby M.J., Barta A., Baxter J.D., Bell G.L., Eberhardt K.L.;
 RT "Analysis of a major human chorionic somatomammotropin gene: Evidence
 for two functional promoter elements.";
 RL J. Biol. Chem. 259:13131-13138(1984).
 RN [2]
 RP SEQUENCE FROM N.A. (GENE CSH2).
 RX MEDLINE=87161235; PubMed=3010560;
 RA Hirt H., Kimmel J., Birnbaum M.J., Chen R.Y., Seeburg P.H.;
 RA Eberhardt K.L., Barta A.;
 RT "The human growth hormone gene locus: structure, evolution, and
 allelic variations.";
 RL DNA 6:59-70(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83160916; PubMed=6300255;
 RA Barrera-Saldana H.A., Seeburg P.H., Saunders G.F.;
 RT "Two structurally different genes produce the same secreted human
 placental lactogen hormone.";
 RL J. Biol. Chem. 258:3787-3793(1983).
 RN [4]
 RP SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).
 RX MEDLINE=89307277; PubMed=2744760;
 RA Choe E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gellius R.E.,
 RA Seeburg P.H.;
 RT "The human growth hormone locus: nucleotide sequence, biology, and
 evolution.";
 RL Genomics 4:479-497(1989).
 RN [5]
 RP SEQUENCE.
 RX MEDLINE=83182010; PubMed=7159009;
 RA Seeburg P.H.;
 RT "The human growth hormone gene family: nucleotide sequences show
 recent divergence and predict a new polypeptide hormone.";
 RL DNA 1:239-245(1982).
 RN [6]
 RP SEQUENCE FROM N.A.
 CC TISSUE=Placenta, and Uterus;
 RX MEDLINE=22382357; PubMed=12477932;
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Berger J.G.,
 RA Klausner R.D., Collins F.S., Wagner T., Shenmen S.M., Scheller G.D.,
 RA Altschul S.F., Zeeberg B., Bietow K.H., Schaefer G.F., Blatt R.K.,
 RA Hopkins R.P., Jordan H., Moore T., Wax S., Wang J., White P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Abert M.,
 RA Stapleton M., Soares M.B., Bonaldi V.P., Casavant T.L., Schaefer G.F.,
 RA Brownstein M.J., Udgin T.R., Toshchinskiy S., Cantarel D., Prange O.,
 RA Raha S.S., Loquellane N.A., Peters G.J., Andrews S.P., Hutchins S.B.,
 RA Bosak S.A., McKernan K.J., Matek A.A., Gaborath P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Guy E.J., Hilly S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettner M., Madan A., Kodanius R., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shvachenko Y., Boultard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 50-217 FROM N.A.
 RX MEDLINE=78071761; PubMed=593368;
 RA Shine J., Seeburg P.H., Martell J.A., Baxter J.D., Goodman H.M.;

PT *Construction and analysis of recombinant DNA for human chorionic
 somatomammotropin.";
 RT Nature 270:494-499(1977).
 RN [8]
 RP SEQUENCE OF 27-217.
 RX MEDLINE=77320197; PubMed=4712450;
 RA Li C.H., Dixon J.S., Cheng D.;
 RT "Amino acid sequence of human chorionic somatomammotropin.";
 RL Arch. Biochem. Biophys. 155:95-110(1973).
 RN [9]
 RP SEQUENCE OF 27-117.
 RX MEDLINE=72016133; PubMed=5286363;
 RA Sherwood L.M., Handwerker S., McLaurin W.D., Lanner M.;
 RT "Amino acid sequence of human placental lactogen.";
 RL Nature New Biol. 233:59-61(1971).
 RN [10]
 RP ERRATUM.
 RA Sherwood L.M., Handwerker S., McLaurin W.D., Lanner M.;
 RL Nature New Biol. 235:54-64(1972).
 RN [11]
 RP INTERCHAIN DISULFIDE BONDS.
 RX MEDLINE=79173081; PubMed=438159;
 RA Schneider A.R., Kowalski K., Russell J., Sherwood L.M.;
 RT "Identification of the interchain disulfide bonds of dimeric human
 placental lactogen.";
 RL J. Biol. Chem. 254:3782-3787(1979).
 CC !- FUNCTION: SIMILAR TO THAT OF SOMATOTROPIN.
 CC !- SUBCELLULAR LOCATION: Secreted.
 CC !- MISCELLANEOUS: THE SEQUENCE OF CSH1 IS SHOWN.
 CC !- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 EMBL: V00573; CAA23836.1;
 EMBL: J00289; AAA98747.1;
 EMBL: K02401; AAA52115.1;
 EMBL: M25894; AAA52116.1;
 EMBL: J03071; AAA52551.1;
 EMBL: J00118; AAA38621.1;
 EMBL: R002717; AAH02717.1;
 EMBL: BC005921; AAH05921.1;
 EMBL: BC020756; AAH20756.1;
 PIR: A26449; A26449.
 PIR: C32435; ICHUG.
 HSP: P01241; 1A22.
 Genew: HGNC:2440; CSH1.
 Genew: HGNC:2441; CSH2.
 MIM: 150200;
 GO: GO:0007565; P:pregnancy; TAS.
 InterPro: IPR001400; Somatomotropin.
 Pfam: PF00103; hormone; 1.
 ProSITE: PS0266; SOMATOTROPIN_1; 1.
 ProSITE: PS00338; SOMATOTROPIN_2; 1.
 Hormone: Placenta; Multigene family; Signal.
 SIGNAL: 1 26
 CHAIN: 27 217 LACTOGEN.
 DISULFID: 79 191
 DISULFID: 208 215
 DISULFID: 208 208 INTERCHAIN (WITH C-215 IN A DIMER).
 DISULFID: 215 215 INTERCHAIN (WITH C-208 IN A DIMER).
 VARIANT: 3 3 P -> A (IN CSH2).
 VARIANT: 104 105 /FTID-VAR_007166.
 VARIANT: 84 84 IS -> L (IN CSH2).
 CONFLICT: 84 84 I -> T (IN REF. 9).
 CONFLICT: 95 95 MISSING (IN REF. 9).
 CONFLICT: 116 116 MISSING (IN REF. 9).

```

FI CONFLICT 134 135 SDD -> BAS (IN REF. 5)
SQ SEQUENCE 217 AA; 25020 MW; 235B0DC7A712F41 CRC64;

Query Match: 47.8%; Score 381; DB 1; Length 217;
Best Local Similarity 82.0%; Pred. No. 4.3e-23;
Matches 61; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
DB 1 TPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
QY 29 TVPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
DB 1 TVPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
QY 64 NNEETQOKSNLELLRISLLIQSWLEPVQ 92
DB 1 NNEETQOKSNLELLRISLLIQSWLEPVQ 92
QY 89 NNEETQOKSNLELLRISLLIQSWLEPVQ 117
DB 1 NNEETQOKSNLELLRISLLIQSWLEPVQ 117

RESULT 10
SOMA_MESAU
ID SOMA_MESAU STANDARD; PRT; 216 AA.
AC P37886;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1 OR GH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Cricetidae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=92063650; PubMed=954881;
RA Southard J.N., Sanchez-Jimenez P., Campbell G.F., Talamantes F.
RT "Sequence and expression of hamster prolactin and growth hormone
messenger RNAs."
RL Endocrinology 129:2965-2971(1991).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This Swiss-Prot entry is copyright. It is prepared through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its context is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S66299; AAB20368.1;
DR PIR: B49159; B49159.
DR HSP: P01246; 1B5T.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA; 24590 MW; 3969CE32A5B51166 CRC64;

Query Match 39.0%; Score 310.5; DB 1; Length 216;
Best Local Similarity 67.0%; Pred. No. 2.5e-23;
Matches 61; Conservative 10; Mismatches 16; Indels 1; Gaps 1.
QY 2 FPTPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
DB 1 FPTPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61

```

```

DB 27 FPAMPPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 85
QY 62 PSNREETOOKSNLELLRISLLIQSWLEPVQ 92
DB 1 PSNREETOOKSNLELLRISLLIQSWLEPVQ 92
QY 86 PTGKEAQORSQSMELLRPSLLIQSWLGPVQ 116
DB 1 PTGKEAQORSQSMELLRPSLLIQSWLGPVQ 116

RESULT 11
SOMA_BALBO
ID SOMA_BALBO STANDARD; PRT; 190 AA.
AC P33092;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin (Growth hormone).
GN GH1.
OS Balaenoptera borealis (Sei whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9768;
RN [1]
SEQUENCE.
RX MEDLINE=8300569; PubMed=7115813;
RA Yudaev N.A., Pankov V.A., Bulatov A.A., Osipova T.A.;
RT "Amino acid sequence of sei whale somatotropin."
RL Biokhimiia 47:1059-1069(1982).
RN [2]
PRELIMINARY PARTIAL SEQUENCE.
RA Osipova T.A., Bulatov A.A., Pankov V.A.;
RT "Structural studies of tryptic peptides from large cyanogen bromide
fragments of sei whale (Balaenoptera borealis) somatotropin."
RL Biokhimiia 47:1589-1599(1978).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC PIR: P01240; PNO140.
DR HSP: P01241; 1AXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary.
FT DISULFID 52 163 BY SIMILARITY.
FT DISULFID 180 188 BY SIMILARITY.
SQ SEQUENCE 190 AA; 21835 MW; 09FBFF6DB14A75D6 CRC64;

Query Match 38.6%; Score 307.5; DB 1; Length 190;
Best Local Similarity 67.0%; Pred. No. 4.3e-23;
Matches 61; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
QY 2 FPTPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
DB 1 FPTPLSLRFDNALRAHRLHOLADTYQFEFAYIPKQKYSFLQNPQTSLSFSSESIPT 61
QY 62 PSNREETOOKSNLELLRISLLIQSWLEPVQ 92
DB 1 PSNREETOOKSNLELLRISLLIQSWLEPVQ 92
QY 60 PANDKDAQORSQSMELLRPSLLIQSWLGPVQ 90
DB 1 PANDKDAQORSQSMELLRPSLLIQSWLGPVQ 90

RESULT 12
SOMA_HORSE
ID SOMA_HORSE STANDARD; PRT; 216 AA.
AC P01245;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=94266171; PubMed=8226352;
RA Ascacio-Martinez J.A., Barrera-Salazar H.A.
RT "Sequence of a cDNA encoding horse growth hormone."
RL Gene 143:299-300(1994).
RN [2]
RP SEQUENCE OF 27-216.
RX MEDLINE=77005410; PubMed=965151;
RA Zakin M.M., Poskus E., Lantton A.A., Ferrara J., Santome J.A.,
RA Bellacha J.M., Paladini A.C.
RT "Primary structure of equine growth hormone."
RL Int. J. Pept. Protein Res. 3:435-444(1976).
RN [3]
RP PRELIMINARY SEQUENCE OF 27-216.
RX MEDLINE=74020362; PubMed=4747845;
RA Zakin M.M., Poskus E., Bellacha J.M., Paladini A.C., Santome J.A.,
RT "The amino acid sequence of equine growth hormone."
RL FEBS Lett. 34:353-355(1973).
RN [4]
RP SEQUENCE OF 69-95 AND 183-216.
RA Zakin M.M., Poskus E., Bellacha J.M., Paladini A.C., Santome J.A.,
RI "Amino acid sequences around the cystine residues in equine growth hormone."
RL FEBS Lett. 25:77-82(1972).
RN [5]
RP SEQUENCE OF 202-216.
RX MEDLINE=68368360; PubMed=4676100;
RA Oliver L., Hartree A.S.
RI "Amino acid sequences around the cystine residues in horse growth hormone."
RL Biochem. J. 109:19-24(1968)
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
CC EM36; 002925; AAA21027.1;
DR HSP; P01246; IBS1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 159
FT DISULFID 205 214
FT SEQUENCE 216 AA; 37AB37383401AC CRC64;
Query Match: 38.5%; Score 306.5; DB 1; Length 216;
Best local similarity 65.2%; Pred. No. 6, 2e-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPTPLSLRFDNKLRAHRLHQLAFDTYQEFEEYIPKQKYSFLONPQTSLSFSISPT 61
D6 27 FPAMPUSLFFANAVLRAQHLDHQAADTYKEFEYIPEGRYS-IQNAAFCFSETIPA 85
QY 62 PSNREETQOKSNLELRISILLIQSNLEPVOL 93
D6 86 FTGKDEAQRSDMELLRFSLLIQSNLGPVOL 117
RESID 13
SOMA_GA:SE STANDARD; PRI: 217 AA.
AC CCKAL;
RT 28-FEB-2003 (Rel. 4); Created
RT 28-FEB-2003 (Rel. 4); Last sequence update
RT 28-FEB-2003 (Rel. 4); Last annotation update
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Galago senegalensis (Northern lesser bushbaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Galago.
OX NCBI_TaxID=9465;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11141192;
RA Adkins R.M., Nekrutenko A., Li W.-H.;
RT "Bushbaby growth hormone is much more similar to nonprimate growth hormones than to rhesus monkey and human growth hormones."
RL Mol. Biol. Evol. 18:55-61(2001).
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
CC EM36; AF292938; AAC44952.1;
DR HSP; P01246; IBS1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 190 BY SIMILARITY.
FT DISULFID 207 215 BY SIMILARITY.
FT SEQUENCE 217 AA; 2F61C031136F05 CRC64;
Query Match: 38.3%; Score 306.5; DB 1; Length 217;
Best local similarity 55.2%; Pred. No. 6, 2e-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPTPLSLRFDNKLRAHRLHQLAFDTYQEFEEYIPKQKYSFLONPQTSLSFSISPT 61
D6 27 FPAMPUSLFFANAVLRAQHLDHQAADTYKEFEYIPEGRYS-IQNAAFCFSETIPA 86
QY 62 PSNREETQOKSNLELRISILLIQSNLEPVOL 93
D6 87 FTGKDEAQRSDMELLRFSLLIQSNLGPVOL 118

```

RESULT 14
SOMA_NYCPY
ID SOMA_NYCEV STANDARD; PRT: 217 AA.
AC QGMB2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Nycticebus pygmaeus (Pygmy slow loris).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirrhini; Lorisae; Nycticebus.
OX NCBI_TaxID=101278;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Wallis O.C., Zhang Y.P., Wallis M.;
RT "Cloning and characterisation of the gene encoding slow loris growth
  hormone.";
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: AJ237562; CAC03504.1; -
DR HSSP: P01246; 1BST.
DR InterPro: IPRO01400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 217 SOMATOTROPIN.
FT DISULFID 79 190 BY SIMILARITY.
FT DISULFID 207 216 BY SIMILARITY.
SQ SEQUENCE 217 AA; 24595 MW; 7F590E77E5D0065F5 CRC64.

Query Match 38.5%; Score 306.5; DB 1; Length 217;
Best Local Similarity 65.2%; Pred. No. 9.7e-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

OY 2 FTPTPLSLRFDNAVLRAERHQLAFTDYQFEAYIPKEQYSHQNTQTSLSFSSEIPT 61
DB 28 FPMPLSSLSFNAVLRAGHLHQLAFTDYQFEAYIPKEQYSHQNTQTSLSFSSEIPT 61
OY 62 PSNREETOQKSNLELLRISLLIOSWLEPVQL 93
DB 67 PTGKDEAQQRSDMELLRFSLLLIOSWLEPVQL 118

RESULT 15
SOMA_MOUSE
ID SOMA_MOUSE STANDARD; PRT: 216 AA.
AC P06880;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1 OR GH.

```

```

Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC MEDLINE=85261358; PubMed=2951252;
RA Linzer D.I.H., Talamantes F.;
RT "Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
  expression of these mRNAs during pregnancy.";
RL J. Biol. Chem. 260:9574-9579(1985).
RN [2];
RP SEQUENCE FROM N.A.
RC SIRAIN-FZTDJ; TISSUE=Liver;
RA Das P., Meyer L., Seyfert H.-M., Brockmann G., Schwerin M.;
RT "Structure of the growth hormone-encoding gene and its promoter in
  mice.";
RL Gene 169:209-213(1996).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: X02891; CAA26650.1; -
DR EMBL: X02891; CAA26650.1; -
DR F01246; 1BST.
DR HSSP: P01246; 1BST.
DR MGI: MGI:95707; Gh.
DR InterPro: IPRO01400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA; 24715 MW; 98666A3AE25D65FC CRC64.

Query Match 38.2%; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 9.7e-23;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

OY 2 FTPTPLSLRFDNAVLRAERHQLAFTDYQFEAYIPKEQYSHQNTQTSLSFSSEIPT 61
DB 27 FPMPLSSLSFNAVLRAGHLHQLAFTDYQFEAYIPKEQYSHQNTQTSLSFSSEIPT 61
OY 62 PSNREETOQKSNLELLRISLLIOSWLEPVQ 92
DB 86 PTGKDEAQQRSDMELLRFSLLLIOSWLEPVQ 116

Search completed: September 16, 2003, 12:38:52
Job time : 12.6732 secs

```

Result No.	Score	Query Length	DB ID	Descriptor
1	436	54.7	217	6 Q8WNE0
2	427.5	52.6	245	4 Q14644
3	407.5	51.1	217	6 Q67469
4	396	49.7	212	6 Q07368
5	396	49.7	217	6 Q07367
6	381	47.8	217	4 Q14407
7	348	43.7	217	6 Q8WNU9
8	341	42.8	202	4 Q14543
9	322.5	40.5	217	6 Q8M174
10	306.5	38.5	216	11 Q70615
11	301.5	37.8	216	6 Q8M173
12	301.5	37.6	216	6 Q83345
13	298.5	37.5	216	11 Q95203
14	297.5	37.3	204	6 Q95205
15	297.5	37.3	216	11 Q95K94
16	297	37.2	217	6 Q8M175


```

OS Macaca mulatta (Rhesus Macaque).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Corropithecoidae:
OC Cercopithecoidea: Macaca.
OC NCBI_TaxID=9544;
RN [1]
RP
RC SEQUENCE FROM N.A.
RC TISSUE=Midpregnancy placenta.
RX MEDLINE=94008724; PubMed=9404617;
RX Glos T.G.; Durrain Y., Fisher J.M., Fowler P.J.;
RT *Cloning of four growth hormone/chorionic somatomammotropin-related
RT complementary deoxyribonucleic acids differentially expressed during
RT pregnancy in the rhesus monkey placenta";
RL Endocrinology 133:1744-1752(1993);
DR EMBL: L16552; AAA1839.12;
DR HSSP: PC1241; IAX;
DR InterPro: IPRC01400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00338; SOMATOTROPIN_2;
SQ SEQUENCE 217 AA: 24542 MW: 255449.5231P2RC GRC64:
Query Match 49.74; Score 366; ES 6; Length 217;
Best Local Similarity 63.24; Pred. No. 2,7,-34;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 3 PTPLSLRFNANMLRAHLHQLAFDTQVEEAYIPKQKSYFIONQTSLSSSESIFIP 62
DB 28 PSVPLSLFPHAMIAHRLHQLAFDTQVEEAYIPKQKSYFIONQTSLSSSESIFIP 67
QY 63 SNREETQOKSNLELR:SLLIQSNLEPVQ 92
DB 88 SNLEETQOKSNLELR:SLLIQSNLEPVQ 117
RESULT 6
Q14407 PRELIMINARY: PRT: 217 AA.
AC Q14407;
DT 01-NOV-1996 (TRENDArel. 01, created)
DT 01-NOV-1996 (TRENDArel. 01, last sequence update)
DT 01-MAR-2003 (TRENDArel. 23, last annotation update)
DE Chorionic somatomammotropin CS-2 (Chorionic somatomammotropin hormone
DE 2).
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
OC NCBI_TaxID=9606;
RN [1]
RP
RC SEQUENCE FROM N.A.
RC MEDLINE=8930727; PubMed=274769;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gellera R.F.,
RA Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
RT evolution.";
RL Genomics 4:479-497(1982).
RN [2]
RP
RC SEQUENCE FROM N.A.
RC MEDLINE=9102558; PubMed=1980158;
RA Vencak-Jones C.L., Phillips J.A.;
RT "Hot spots for growth hormone gene deletions in homologous regions
RT outside of Alu repeats.";
RL Science 250:1745-1748(1990).
RN [3]
RP
RC SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/CCDSJ databases.
RN [4]
RP
RC SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/CCDSJ Databases.

```

```

CY      63 SNKRETOQSKNSLELRIISLLLIQSWSLPVQLGTGPRFVNOHL 100
       ::::  ::::|::|::|::|::|::|::|::|::|::|::|::|::|
DL      88 FKKEEMLGKSXVELIHIISLLLIQSWSLEPMQ-RLGISIFANSOL 128
                                         |::|::|::|::|::|::|

RESULT 10
QSM173
ID    Q70615 PRELIMINARY: PRT: 216 AA.
AC    Q70615;
DC    C1-AUG-1998 (TRENBLrel_07, Created)
DT    C1-AUG-1998 (TRENBLrel_07, Last sequence update)
DE    C1-WAR-2003 (TRENBLrel_23, Last annotation update)
DS    Growth hormone precursor.
OS    Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sp
OC    Nannospalax.
CN    NCBI_TAXID=30637;
LN    [1]
PE    SEQUENCE FROM N.A.
RP    MEDLINE=99124645; PubMed=9924177;
RX    Lioupis A., Nevo E., Wallis M.;
RT    "Cloning and characterisation of the gene encoding mole
RT    ehrenbergi growth hormone.";
PL    J. Mol. Endocrinol. 22:29-36(1999).
DR    EMBL: AF025819; CAA06716.1; -.
DR    HSS: POL241; IAXI.
DR    InterPro: IPRO014C0; Somatotropin.
DR    Pfam: PF03163; hormone; 1.
DR    PRINTS: PR00836; SOMATOTROPIN.
DR    PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR    PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW    Signal.
FT    SIGNAL.
FT    CHAIN.
SQ    SEQUENCE 216 AA; 24637 MW; EEAB9A523BACADF: CRC64;

Query Match          39.5%; Score 306.5; DB 11; Length
Best local similarity 65.9%; Pred. No. Id-24;
Matches 69; Conservative 13; Mismatches 17; Indels 1

CY      2 FETPILSLDFDNAMLRHRLHQALFDITYOEFEFFAYIPKEQKYSLOF 100
       ::::  ::::|::|::|::|::|::|::|::|::|::|::|::|::|
DL      27 FPAPMLSNLFANAVLRQHILQSLANDTYKEFERAYIPEGQRYS-IO 128
                                         |::|::|::|::|::|::|

CY      62 PSNRETQSKNSLEELRISLLIQSWLEPVMQ 92
       ::::  ::::|::|::|::|::|::|::|::|::|::|::|::|::|
DL      86 PTGREAEACRSDMELRFESLLLIQSWSLGPVQ 116
                                         |::|::|::|::|::|::|

RESULT 11
QSM173
ID    Q5M173 PRELIMINARY: PRT: 216 AA.
AC    Q5M173;
DC    01-OCT-2002 (TRENBLrel_22, Created)
DT    01-OCT-2002 (TRENBLrel_22, Last sequence update)
DE    01-WAR-2003 (TRENBLrel_23, Last annotation update)
DS    Growth hormone precursor.
OS    Delphinus delphis (Saddleback dolphin) (Black sea dolphin).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
OC    Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti
OC    Delphinus.
CN    NCBI_TAXID=9728;
LN    [1]
PE    SEQUENCE FROM N.A.
RP    TISSUE=Liver.
RX    Manion Z., Wallis O.C., Wallis M.;
RT    "Cloning and characterisation of the GH gene from the com
RT    (Delphinus delphis).";
DR    Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR    EMBL: AJ492191; CAD37292.1; -.

```



```

DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00148; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 216 GROWTH HORMONE.
SQ SEQUENCE 216 AA; 24509 MW; 1E3467A44CCE502 CRC64;

Query Match 37.9%; Score 301.5; DB 6; Length 216;
Best Local Similarity 64.8%; Pred. No. 3.4e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FTTPIPSRLFDNAMLRAHLRHLQALADTYQEFEEAYIPKEQKYSFIQNPTCSLSFSSESIPT 61
DB 15 FAMSLSRLFANAVLRAQHLHQLADTYKFEFRTYIPEQQRYS-IQNTQVAFCSFTIPA 73
QY 62 PSNREETQOKSNDELLRISLLLIQSWLGPVQ 92
DB 74 PTKSNEAQKSDLELLRISLLLIQSWLGPVQ 104

RESULT 12
Q8HYE5 PRELIMINARY: PRI: 216 AA.
AC Q8HYE5;
DI 01-MAR-2003 (TrEMBLrel. 23, Created);
DI 01-MAR-2003 (TrEMBLrel. 23, Last sequence update);
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Growth hormone precursor.
OS Alluropoda melanoleuca (Giant panda).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Alluropoda.
CX NCBI_TaxID=9645;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Liao M., Zhu M., Zhang A.;
RT "Cloning and expression of cDNA encoding growth hormone from
Alluropoda melanoleuca."
RL Submitted (AUG-2002) to the EMBL/GenBank/CCBJ databases.
DR EMBL: AF549336; AAN77228.1; -.
FT SIGNAL 1 26 POTENTIAL.
FT SIGNA;
SQ SEQUENCE 216 AA; 24283 MW; 44EC17E44RC656 CRC64;

Query Match 37.9%; Score 301.5; DB 6; Length 216;
Best Local Similarity 64.8%; Pred. No. 3.4e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FTTPIPSRLFDNAMLRAHLRHLQALADTYQEFEEAYIPKEQKYSFIQNPTCSLSFSSESIPT 61
DB 15 FAMSLSRLFANAVLRAQHLHQLADTYKFEFRTYIPEQQRYS-IQNTQVAFCSFTIPA 73
QY 62 PSNREETQOKSNDELLRISLLLIQSWLGPVQ 92
DB 74 PTKSNEAQKSDLELLRISLLLIQSWLGPVQ 116

RESULT 13
Q9R2C3 PRELIMINARY: PRI: 216 AA.
AC Q9R2C3;
DI 01-MAY-2000 (TrEMBLrel. 13, Created);
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update);
DI 01-MAY-2000 (TrEMBLrel. 23, Last annotation update);
DE Growth hormone.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]

```

```

SEQUENCE FROM N.A.
RA Nguyen T.N.K., Liebhauer S.A.;
RT "Mouse Growth Hormone Locus: Nucleotide Sequence and Phylogenetic
Analyses."
RL Submitted (AUG-1995) to the EMBL/GenBank/CCBJ databases.
DR EMBL: U34362; AAC99988.1; -.
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 216 AA; 24682 MW; FC2A06DAC2536B18 CRC64;

Query Match 37.5%; Score 298.5; DB 11; Length 216;
Best Local Similarity 63.7%; Pred. No. 7.2e-24;
Matches 56; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 2 FTTPIPSRLFDNAMLRAHLRHLQALADTYQEFEEAYIPKEQKYSFIQNPTCSLSFSSESIPT 61
DB 15 FAMSLSRLFANAVLRAQHLHQLADTYKFEFRTYIPEQQRYS-IQNTQVAFCSFTIPA 73
QY 62 PSNREETQOKSNDELLRISLLLIQSWLGPVQ 92
DB 74 PTKSNEAQKSDLELLRISLLLIQSWLGPVQ 116

RESULT 14
Q9R2C5 PRELIMINARY: PRI: 204 AA.
AC Q9R2C5;
DI 01-FEB-1997 (TrEMBLrel. 02, Created);
DI 01-FEB-1997 (TrEMBLrel. 02, Last sequence update);
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Placenta; growth hormone (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
CX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDJINE=97050633; PubMed=8895361;
RA Lacroix M.C., Devinoy E., Servely J.L., Puissant C., Kann G.;
RT "Expression of the growth hormone gene in ovine placenta: detection
and cellular localization of the protein."
RL Endocrinology 137:4886-4892(1996).
DR EMBL: U49063; AAC48679.1; -.
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
FT NON_TER 1 1
SQ SEQUENCE 204 AA; 23462 MW; EBBB9451892635C6 CRC64;

Query Match 37.3%; Score 297.5; DB 6; Length 204;
Best Local Similarity 64.8%; Pred. No. 8.6e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FTTPIPSRLFDNAMLRAHLRHLQALADTYQEFEEAYIPKEQKYSFIQNPTCSLSFSSESIPT 61
DB 15 FAMSLSRLFANAVLRAQHLHQLADTYKFEFRTYIPEQQRYS-IQNTQVAFCSFTIPA 73
QY 62 PSNREETQOKSNDELLRISLLLIQSWLGPVQ 92
DB 74 PTKSNEAQKSDLELLRISLLLIQSWLGPVQ 104

RESULT 15
Q9JMK4

```

